치의과학박사학위논문

Comparison of Individualized Facial Growth Prediction Models Based on the Multivariate Partial Least Squares Method and Artificial Intelligence Developed by TabNet Deep Neural Network

부분최소제곱 방법과 인공지능을 이용한 측모 성장 예측 방법 개발 및 비교 연구

2023년 2월

서울대학교 대학원

치의과학과 치과교정학 전공

문 준 호

Comparison of Individualized Facial Growth Prediction Models Based on the Multivariate Partial Least Squares Method and Artificial Intelligence Developed by TabNet Deep Neural Network

지도교수 이 신 재

이 논문을 치의과학박사 학위논문으로 제출함

2022년 12월

서울대학교 대학원

치의과학과 치과교정학 전공

문 준 호

문준호의 치의과학박사 학위논문으로 인준함

2023 년 1 월

위 원 장 안 석 준 (인)

부위원장 이 신 재 (인)

위 원 양 일 형 (인)

위 원 박 영 석 (인)

위 원 허 수 복 (인)

- Abstract -

Comparison of Individualized Facial Growth Prediction Models Based on the Multivariate Partial Least Squares Method and Artificial Intelligence Developed by TabNet Deep Neural Network

Jun-Ho Moon, BS, DDS

Department of Orthodontics, Graduate School,

Seoul National University

(Directed by Professor ***Shin-Jae Lee****, DDS, MSD, PhD, PhD)*

[Introduction]

Craniofacial growth has long been considered an important topic in the field of clinical orthodontics. Several growth prediction methods have been developed, however, individual variations in growth make prediction challenging. In addition, growth prediction requires predicting numerous highly inter-correlated variables, which limits the use of various statistical techniques.

When significant numbers of input and output variables are highly correlated with each other, such as soft tissue responses after orthognathic surgery, prediction models based on the partial least squares (PLS) method showed better

predictive performance than conventional ordinary least squares (OLS) methods. Therefore, the PLS method might be useful in predicting growth by reflecting individual growth variations and solving correlation issues.

Meanwhile, attempts to apply artificial intelligence (AI) in the field of dentistry have been increasing. Recently, the TabNet algorithm has been developed to apply deep neural networks (DNNs) to tabular data. Applying TabNet algorithm to tabular growth data might enable accurate growth predictions.

The purpose of this study was to develop and compare facial growth prediction models incorporating individual skeletal and soft tissue characteristics based on the PLS method and artificial intelligence.

[Materials and methods]

Serial longitudinal lateral cephalograms were collected from 303 children (166 girls and 137 boys), who had never undergone orthodontic treatment. Growth prediction models were devised by applying the multivariate PLS algorithm and AI developed by TabNet deep neural network, with 161 predictor variables. Response variables comprised 78 lateral cephalogram landmarks. T-tests were performed to compare the prediction accuracy between the two methods. Multiple linear regression analysis was performed to investigate factors influencing growth prediction errors. Confidence ellipses were depicted to investigate the pattern of prediction errors and to evaluate the effect of growth variability on the accuracy of prediction models.

[Results]

Using the leave-one-out cross-validation method, a PLS model with 30 components was developed. For the AI-based prediction model, optimal hyperparameters were selected after hyperparameter tuning. Among the 78 landmarks, the AI-based model was more accurate in 55 landmarks. The PLS method was more accurate in 10 landmarks, including cranial base landmarks, which generally showed less growth variability. The remaining 13 landmarks showed no statistical difference between two methods. When uncertainty was high, it was more advantageous to use AI for growth prediction. On average, the AI-based model showed less prediction error by 1.11mm than the PLS-based model. Younger age at prediction resulted in greater prediction error (0.01 mm per year). In addition, prediction error increased in proportion to the growth prediction interval (0.14 mm per year). Girls, subjects with Class II malocclusion, skeletal landmarks, and landmarks on the maxilla showed more accurate prediction results than boys, subjects with Class I or III malocclusion, soft tissue landmarks, and landmarks on the mandible, respectively.

[Conclusions]

The AI-based model predicted growth more accurately than the PLS-based model. The prediction error of the prediction model was proportional to the remaining growth potential. PLS and AI growth prediction seemed to be a versatile approach that can incorporate large numbers of predictor variables to predict numerous landmarks for an individual subject.

Key Words: Growth prediction, Partial least squares method, Artificial intelligence, Deep learning

Student Number: 2018-38862

CONTENTS

Ⅰ. INTRODUCTION

Craniofacial growth is a fundamental topic in orthodontics. Particularly in clinical practice, growth prediction assists orthodontists in formulating treatment plans and visualizing therapeutic outcomes to accomplish satisfying results for growing patients. Various growth prediction methods have been developed, with respect to both direction and magnitude;[1-17] however, accurate growth prediction remains challenging, due to the extremely variable nature of growth in individuals.

Growth is a complex process affected by both genetic and environmental factors, and varies according to sex and ethnicity.[1,3,4,15,18] Variation in craniofacial growth according to cephalometric characteristics has been reported previously.[7,18-21] Growth prediction methods estimate a patient’s residual growth based on average annual increments, as well as the anticipated amount of growth added to the patient’s current state. As summarized in Table 1, growth prediction methods included specific cephalometric templates and guides, such as mesh diagrams,[9,22] grids,[5,23] templates,[10] and Ricketts’ visual treatment objective (VTO)[11,12,24]. However, these approaches do not account for individual variation; rather, average growth per year is applied to every patient. Subsequent studies have used more sophisticated approaches based on multivariate statistical methods,[14,25] such as Bayesian theorem,[13] a multilevel model,[2] and application of nonlinear growth functions.[6,7,17] Yet, growth prediction remains among the most daunting challenges in orthodontics. Numerous factors, such as innate skeletal and soft tissue variables, as well as a large amount of biological information, such as age and sex, must be considered to produce accurate and clinically applicable predictions.

When considerable numbers of both input predictor variables and output response variables are highly correlated with one another, prediction models based on the partial least squares (PLS) method demonstrated superior predictive performance over conventional ordinary least squares (OLS) methods, such as linear regression models.[26-30] A number of previous reports have demonstrated that the PLS algorithm was significantly more accurate for predicting treatment outcomes than OLS-based methods. The improved prediction capability of the PLS method may be due to its ability to control for significant correlations among the skeletal and soft tissue variables of individual patients.[26-30] Furthermore, post-treatment changes are affected by various factors, including age and sex, among others. As predicting treatment outcomes and growth changes likely involve similar aspects, the PLS method is expected to be a useful tool for predicting growth by considering various factors. Through linear combination of numerous variables via matrix algebra, PLS can reflect the skeletal and soft tissue characteristics of an individual.

Recently, the use of artificial intelligence (AI) in the field of dentistry has gradually attracted attention. In the field of orthodontics, there have been attempts to apply AI in cephalometric landmark detection and orthodontic diagnosis.[31-36] In these cases, one of the deep learning methods, deep neural networks (DNN), are mainly used.[37] A DNN, which is capable of modeling complex nonlinear relationships, is an artificial neural network comprising multiple layers connecting the input and output layers. Lately, the DNN architecture TabNet, developed to apply deep learning to tabular data, was shown to outperform other DNN algorithms on tabular datasets.[38] Therefore, applying TabNet algorithm in growth data, which can be considered as tabular, might show promising growth prediction results.

The purpose of this study was to develop and compare facial growth prediction models based on the PLS method and artificial intelligence.

Ⅱ. REVIEW OF LITERATURE

As shown in Table 1, various growth prediction methods have been developed, from the famous Ricketts' VTO[11,12,24] to using parental data, which was somewhat different from preexisting methods, for predicting offspring's growth.[14] Previous growth prediction methods could be categorized as follows: 1) adding increments to present size, 2) using skeletal maturity to predict mandibular growth, 3) using statistical methods to predict growth.

Adding increments to present size

The methods in the first category add an average increment to the present size to predict growth. These are the most classical and well-known methods. The advantage of these methods is that it is easy to understand and can be applied without difficulty. However, in these methods, individual variations are not considered, and only average growth is applied to every patient.

The mesh diagram is a grid system used for craniofacial growth analysis and prediction.[9,22] It consists of 24 rectangles of the same size drawn on a cephalometric tracing. The middle four rectangles are called core rectangles. The size of the core rectangle varies from person to person and determines the size of the remaining rectangles. In the mesh diagram method, it was expected that the size of the core rectangle would increase by a certain amount and other structures would grow in proportion to the size of the core rectangle.

Ricketts proposed arcial growth of the mandible.[11] In this study, three arcs passing through protuberance menti were presented. Ricketts insisted that when arcial growth of the mandible and annual mandibular growth of 2.5 mm were combined, the size and shape of the mandible can be predicted with considerable accuracy. In addition to the arc-shaped growth of the mandible, Ricketts developed a method called VTO that can predict facial growth and soft tissue changes.[12,24]. Growth prediction was performed by dividing the craniofacial region into six areas: cranial base prediction, mandibular growth prediction, maxillary growth prediction, occlusal plane prediction, the location of dentition, and the soft tissue of the face.

Some studies used forecast grid to predict growth.[5,23] Different templates were used by gender and age. In this growth prediction method, there were forecast grids for several anatomic landmarks. The growth of a particular landmark was anticipated to move from its current position along the axes of the forecast grid. To increase the accuracy of prediction, grid units were set differently by the starting age of prediction, years of prediction, and gender. The disadvantage of this method was that the same amount of growth was assumed regardless of the growth pattern of the face.

In the study by Popovich and Thompson, craniofacial templates were used to predict growth.[10] These templates were provided differently according to gender and vertical growth patterns. Growth patterns were classified into horizontal, vertical, and average, and a total of six craniofacial templates were provided.

Using skeletal maturity to predict mandibular growth

The methods in the second category use skeletal maturity to predict mandibular growth. These methods have primarily focused on predicting mandibular growth potential in Class III subjects, since, clinically, the prediction of growth is necessary for children with a skeletal discrepancy.

In the study by Sato et al., hand-wrist radiographs were used to predict mandibular growth in Japanese girls.[39] In this study, the ossification event method, the growth potential method, the growth percentage method, the growth chart method, and the multiple regression method were presented as prediction models. Among these prediction models, the growth potential method and the growth percentage method were proved to be the most accurate method for predicting mandibular length in Japanese girls. In the growth potential method, the relationship between bone age and growth potential was obtained by simple linear regression. On the other hand, in the growth percentage method, the ratio between the current mandibular length and the final mandibular length was determined by bone age.

The study of Mito et al. used cervical vertebral bone age to predict mandibular growth potential of growing Japanese girls.[40] Simple linear regression analysis was performed with cervical vertebral bone age as an independent variable.

Chen et al. also made similar predictions using cervical vertebrae as an indicator of skeletal maturation.[41] The difference was that multiple linear regression was used, and the dimensions of the third and fourth cervical vertebrae were used as independent variables.

Using statistical methods to predict growth

Finally, the methods in the third category use statistical methods to predict craniofacial growth. Multivariate method to predict craniofacial pattern was proposed.[25] In this study, craniofacial patterns were classified into nine types using factor analysis and cluster analysis.[42] The craniofacial pattern at 9.5 years was used to predict the pattern of 17.5 years.

An attempt was made to predict the craniofacial growth pattern of offspring from the craniofacial form of parents.[14] Using parental information, rather than average growth curves, in predicting growth was proposed, since it was known that there is a high correlation between the offspring's craniofacial form and the parents'. Principal component analysis and cluster analysis were used to classify craniofacial patterns into four types. The offspring's parents were assigned as similar parent or dissimilar parent depending on whether the craniofacial pattern coincided. Multiple linear regression was used to predict five craniofacial variables implying the distance between anatomical landmarks. For each variable, parents' measurements were used as independent variables to predict offspring's measurement at given ages.

In the study by Rudolph et al., multivariate analysis was used to predict the growth of skeletal Class II samples.[13] While previous studies had predicted the amount or direction of craniofacial growth, the study focused on predicting whether growth was favorable or not. Skeletal Class II samples were divided into favorable or unfavorable growth according to the degree of ANB angle improvement from 8 to 18 years of age. Whether a sample would be a good grower or not was predicted by cephalometric measurements and covariance matrix of good growers using Bayesian theorem.

The study of Chvatal et al. used multilevel model to predict craniofacial growth.[2] Longitudinal growth curves for various angular and horizontal cephalometric measurements were developed. Unlike other studies, polynomial terms were included in the growth prediction model.

III. MATERIALS AND METHODS

Growth Data Collection

Subjects comprised 303 growing patients (166 girls and 137 boys), who had not undergone any orthodontic or orthopedic treatment and had at least two serial lateral cephalometric images taken at Seoul National University Dental Hospital, Seoul, Korea, from June 29, 2006 to December 20, 2019. Mean subject ages at the beginning and end of the growth observation period were 10.9 and 14.2 years, respectively (Figure 1). Approximately three-quarters of patients had skeletal Class II or III malocclusion (Table 2), consistent with the proportion of patients with malocclusion visiting the university-affiliated hospital.[43-45]

Although subjects initially wanted to receive active orthodontic treatment at their first visit, treatment did not begin immediately for various reasons. Some subjects had such a severe skeletal discrepancy that observation was necessary until their growth ceased, before they could receive combined surgical-orthodontic treatment. For other subjects, reasons for treatment postponement included finances, poor personal timing, and/or other unreported personal issues.

The institutional review board for the protection of human subjects of the Seoul National University Dental Hospital, Seoul, Korea, reviewed and approved the research protocol (ERI 19007).

Inclusion and Exclusion Criteria

The exclusion criteria were cleft lip and palate, and a syndromic or medically compromised condition. Simple space maintainers were considered to have little impact on growth; therefore, subjects who had used one were included in the present study. For every patient, serial lateral cephalometric radiographs were taken at least twice during the growth observation period. The characteristics of the subjects included in this study are summarized in Table 2.

Cephalometric tracing and landmark identification, both at the beginning (T1) and end (T2) of growth observation, were manually performed for all images by a single examiner (SJL). A total of 46 hard tissue and 32 soft tissue landmarks were identified. To orient consecutive images to the same head position, the horizontal reference plane was set to Sella-Nasion +7°, with its origin at Sella following along the Sella-Nasion plane.[46] The anatomic landmarks, reference planes, and coordinate system used in the study are presented in Figure 2.

Predictor Variables, Response Variables, and Validation

Predictor variables were a heterogeneous set, including individual characteristics (Table 3) that could be categorized into 1) demographic (age and sex); 2) molar relationship; 3) ages before and after the growth observation period; and 4) Cartesian (x,y) coordinates of 78 anatomic landmarks. A total of 161 predictor variables were incorporated into the input X matrix.

Response variables comprised the *x* and *y* directions of 46 hard and 32 soft tissue landmarks after the period of growth observation. A total of 156 response variables were incorporated into the output Y matrix.

When developing a prediction model, a validation process is essential to evaluate the prediction accuracy of subjects that are not used to construct the model. Validation was performed by applying the prediction equation to new data (the *test data*) that were not used in the prediction model building procedure. The resultant test errors (also known as *validation errors*) were computed using the leave-one-out cross-validation technique (LOOCV). Given a total number of subjects, *N*, LOOCV constructs a prediction equation *N* times, with all the subjects except one. Then, the prediction equation is applied to the excluded subject.[47]

Predictor variables, response variables, and the validation process were identical for both growth prediction models.

Predicting Growth with the Partial Least Squares Method

In constructing the prediction model based on the PLS method, this study followed similar procedures to a previous study, where a detailed description of building PLS models were presented.[48] After construction of the prediction model, training errors were calculated on sample data to evaluate the goodness-of-fit of the prediction model.

The optimal number of PLS components, a linear combination of predictor variables including key information of the input X matrix, was determined by comparing test errors (Figure 3). In this study, the root mean squared error of prediction was used because the prediction errors in opposite directions offset each other (Figure 3).[49,50]

Predicting growth with Artificial Intelligence TabNet Algorithm

As previously stated, the AI algorithm used in this study was TabNet, a DNN architecture. The TabNet model was modified to build the growth prediction model by using the Python programming language (Python Software Foundation, Wilmington, Delaware, USA).

Neural networks are known to be parameterized by various hyperparameters. These include epoch, patience, learning rate, batch size, etc. Adequate values for each hyperparameter are known to be important for the model to function properly.[51] Among the hyperparameters, e*poch* refers to a one complete pass of the training dataset through the algorithm. Early stopping condition of the training process, when the performance of the model no longer improves with training, is parametrized by hyperparameter *patience.* The Synthetic Minority Oversampling Technique (SMOTE), one of the oversampling methods, was used to deal with the class imbalance problem in the data.[52] The hyperparameter *SMOTE* value refers to the amount of oversampling applied to the raw data.

Hyperparameter tuning, the process of finding optimal hyperparameter values, was performed by comparing the performance of models with various epoch, patience, and SMOTE values. After hyperparameter tuning, the final prediction model was constructed.

Analyzing Growth Prediction Accuracy and Comparing Prediction Models

All of the PLS and AI models were constructed and tested built on a single desktop computer with recommended specifications for modeling process. Its CPU processor that was critical for the PLS model building time was Intel Core i9-12900K. The GPU graphic card that was crucial for constructing deep learning models was NVIDIA GeForce RTX 3090 Ti. The Linux desktop was used on Ubuntu version 22.04 LTS of Linux distribution.

After establishing the final growth prediction models using the PLS method and the TabNet AI algorithm, predictions were made on the test data using LOOCV to evaluate and compare the predictive performance. The difference between predicted and actual growth for 156 response variables, the x and y coordinates of 46 skeletal and 32 soft tissue landmarks, was calculated for each method.

When a one-dimensional evaluation was performed, larger or smaller predictions compared to actual values yielded positive and negative errors which offset each other. Therefore, the Euclidean distance between actual growth and prediction result of a given landmark was calculated. T-tests with Bonferroni correction was used to compare the prediction accuracy between the PLS method and the

TabNet AI algorithm.

Multiple linear regression analysis, with the absolute value of the prediction error as a dependent variable, was performed to investigate the influence of individual characteristics and landmark attributes on the accuracy of the growth prediction model.

For a two-dimensional evaluation, scatterplots with 95% confidence ellipses were drawn to visualize the pattern of prediction errors.

To evaluate the effect of variability and pattern of growth on the accuracy of prediction models, 95% confidence ellipses were depicted.

The open source statistics program, Language R,[53] was used.

Ⅳ. RESULTS

For the PLS growth prediction model, the optimal model was selected, based on the root mean squared error of a prediction curve (Figure 3). As the number of PLS components increased, test errors initially decreased, but gradually increased as the maximum number of components was reached. Consequently, in this study, the optimal prediction model chosen included 30 PLS components. Figure 4 shows the training and test errors, in the form of mean absolute errors, for several selected anatomical landmarks. Similar patterns were observed for the training and test errors. The magnitude of errors and the differences between the training and test errors tended to increase as landmarks were located at more inferior parts of the face.

The optimal hyperparameters were selected by graphically comparing AI-based growth prediction models with various hyperparameter values (Figure 5). In addition, a linear model was used secondarily in hyperparameter tuning. As a result, in this study, epoch 1,000,000, patience 30,000, and SMOTE 0.1 were selected as the optimal hyperparameters.

The comparison results between prediction models using t-tests are summarized in Table 4. Of the 78 anatomical landmarks, the AI-based prediction model showed better prediction accuracy in 55 landmarks. The PLS-based prediction model was more accurate in 10 landmarks. There was no statistical difference in the remaining 13 landmarks.

As shown in Table 3, the results from multiple linear regression analysis indicated that, on average, the prediction error of the AI-based model was about 1.11 mm smaller than that of the PLS-based model. In addition, the prediction error increased in proportion to the growth prediction interval (0.14 mm per year). Further, prediction error was greater with younger age at prediction (0.01 mm per year). Conversely, the older the age at the prediction, the more accurate the prediction results.

Girls, subjects with Class II malocclusion, skeletal landmarks, and landmarks on the maxilla had lower prediction errors than boys, subjects with Class I or III malocclusion, soft tissue landmarks, and landmarks on the mandible, respectively. There was no statistical difference in prediction errors according to the direction of growth (Table 3).

The two-dimensional patterns of growth prediction errors for representative landmarks are shown from figure 6, A to figure 6, S. Of the landmarks shown, only the posterior nasal spine showed no statistical difference between the two prediction methods. The PLS method showed better prediction accuracy in cranial base landmarks. The prediction results using the AI method were statistically more accurate for the remaining landmarks.

Confidence ellipses indicating the variability and pattern of growth of each landmark are shown in figure 7. To increase the clarity of the graph, some of the landmarks where AI-based prediction was more accurate in t-tests were omitted.

Figure 8 illustrates real case comparisons between actual growth and prediction results. To generate a smooth curve for the soft tissue profile line, cephalometric landmarks were connected using spline functions. The prediction results were far from perfect, but varied among subjects. Overall, AI-based predictions seemed a little closer to actual growth.

V. DISCUSSION

The primary purpose of this study was to develop an automated and reliable growth prediction model that can reflect individual characteristics. Craniofacial growth is considered complex and difficult to predict, since it is influenced by various factors, including sex, ethnicity, and morphological characteristics, among others. To predict such complex skeletal and soft tissue changes accompanied by growth, the present study applied the PLS method, which is capable of reflecting a vast number of predictor variables, and of predicting numerous soft and hard tissue landmarks in an individual subject. This study also attempted to overcome many challenges in predicting growth by using TabNet, a state-of-the-art deep learning algorithm.

On the whole, the AI model predicted growth more accurately than the PLS model. According to multiple linear regression analysis, on average, the prediction error of AI-based model was 1.11 mm smaller than that of PLS-based model. The growth prediction accuracy of the prediction models was different according to the landmarks to be predicted. The results from t-tests indicated that, among the cephalometric landmarks used in this study, the AI model was more accurate in 71% (55 out of 78) of the landmarks. The PLS-based prediction showed higher accuracy in 10 landmarks, mainly including cranial base landmarks such as Nasion, Porion, Orbitale, and Basion. Regarding the growth variability of each landmark (Figure 7), the PLS method showed better performance in landmarks with small variations. Conversely, landmarks where AI was more accurate generally showed great variability in growth. In other words, AI was powerful when uncertainty was high. This tendency might be useful in choosing which method to use in building prediction models. To support this hypothesis, other prediction models, such as predicting soft tissue responses after orthognathic surgery or orthodontic treatment, are worth further research.

From the clinical perspective, the test error represents the criteria for prediction accuracy, while the training error may reflect the goodness-of-fit of the model. The results demonstrate that the test errors of the prediction model tended to increase with landmarks located in more inferior positions. The reason for the low predictive accuracy of landmarks located in the more inferior portion of the face may be the distance from the cranial base. The prediction results for anatomical landmarks located on the mandible were less accurate than those for landmarks on the maxilla (Figures 4 and 8).

The growth prediction error was greater in boys with Class III malocclusion than in girls with Class II malocclusion. We speculate that this may be because, if other conditions, such as age at prediction and growth observation period, were the same, then boys with Class III malocclusion would have greater residual growth potential than girls with Class II malocclusion.

Prediction results were less accurate for soft tissue. We speculate that soft tissues changes did not follow those of hard tissue in a one-to-one manner. Further, soft tissue landmarks may have been affected by varying subject posture.

The construction of prediction equations takes significant amount of time since model building procedure requires multiple iterations and training process for PLS and AI algorithms, respectively. In addition, applying the LOOCV technique as a validation method takes much longer than applying any other type of validation method.[26,47] Consequently, the time spent in establishing prediction equations was more than 10 days for the AI-based model, while about 20 minutes for the PLS-based model; however, once the prediction model was built, the time to produce a prediction result was only a few milliseconds. This is because, unlike the model building procedures, the prediction process involved simple and fast computations without implementing complicated iteration or training procedures. Regarding the large difference in model construction time, it may seem reasonable that the prediction results of the AI-based model were more accurate. Nevertheless, computer-aided clinical environments would be an essential condition for practical application of this growth prediction model.

The current study applied advanced statistical and deep learning approaches; however, growth prediction performance was not as accurate as envisaged. Although imperfect and inaccurate, the prediction model presented here (see the real case application shown in Figure 8) may be useful as a rough guide, which is better than having no means of estimating growth changes – especially when used alongside other digitally-derived methods by providing automated and rapid results.

A strength of the present study is that the data included a larger number of patients and more input and output variables than previous growth prediction studies (Table 1). A limitation of the current study is that the growth observation period varied among patients (Figure 1). The way that growth is interpreted may vary according to the measurement method applied and the observation interval.[6,7] In the present study, growth observation intervals were not prearranged. Rather, subjects who had undergone serial cephalograms were collected retrospectively through medical record collation. Consequently, the interval for growth observation ranged from 1.0 to 13.2 years. Another limitation is that the growth prediction model could not consider the effect of age-related differential growth. Inclusion of additional variables that reflect skeletal age may be necessary.

Ⅵ. CONCLUSIONS

1. The PLS and AI growth prediction models presented here are versatile and incorporates a large number of predictor variables, as well as predicting numerous landmarks in individual subjects.

2. In general, the TabNet AI algorithm predicted growth more accurately than the PLS method. However, the PLS method was favorable in predicting landmarks with low variability.

3. Further refinement using nonlinear age covariates and additional variables reflecting skeletal age may result in a more accurate prediction formula.

ACKNOWLEDGMENT

A part of the data in this PhD dissertation was published in the November, 2022, issue of the Angle Orthodontist (Volume 92, pages 705-713).

REFERENCES

1. Hersberger-Zurfluh MA, Papageorgiou SN, Motro M, Kantarci A, Will LA, Eliades T. Genetic and environmental components of vertical growth in mono- and dizygotic twins up to 15-18 years of age. *Angle Orthod*. 2021;91:384-390.
2. Chvatal BA, Behrents RG, Ceen RF, Buschang PH. Development and testing of multilevel models for longitudinal craniofacial growth prediction. *Am J Orthod Dentofacial Orthop*. 2005;128:45-56.
3. Desai RJ, Iwasaki LR, Kim SM, Liu H, Liu Y, Nickel JC. A theoretical analysis of longitudinal temporomandibular joint compressive stresses and mandibular growth. *Angle Orthod*. 2022;92:11-17.
4. Hennessy RJ, Stringer CB. Geometric morphometric study of the regional variation of modern human craniofacial form. *Am J Phys Anthropol*. 2002;117:37-48.
5. Johnston LE. A simplified approach to prediction. *Am J Orthod*. 1975;67:253-257.
6. Lee SJ, An H, Ahn SJ, Kim YH, Pak S, Lee JW. Early stature prediction method using stature growth parameters. *Ann Hum Biol*. 2008;35:509-517.
7. Lee YS, Lee SJ, An H, Donatelli RE, Kim SH. Do class iii patients have a different growth spurt than the general population? *Am J Orthod Dentofacial Orthop*. 2012;142:679-689.
8. Lundstrom A, Woodside DG. Individual variation in growth directions expressed at the chin and the midface. *Eur J Orthod*. 1980;2:65-79.
9. Moorrees CF, Lebret L. The mesh diagram and cephalometrics. *Angle Orthod*. 1962;32:214-231.

10. Popovich F, Thompson GW. Craniofacial templates for orthodontic case analysis. *Am J Orthod*. 1977;71:406-420.

11. Ricketts RM. A principle of arcial growth of the mandible. *Angle Orthod*. 1972;42:368-386.

12. Ricketts RM. The value of cephalometrics and computerized technology. *Angle Orthod*. 1972;42:179-199.

13. Rudolph DJ, White SE, Sinclair PM. Multivariate prediction of skeletal class ii growth. *Am J Orthod Dentofacial Orthop*. 1998;114:283-291.

14. Suzuki A, Takahama Y. Parental data used to predict growth of craniofacial form. *Am J Orthod Dentofacial Orthop*. 1991;99:107-121.

15. Ursi WJ, Trotman CA, McNamara JA, Jr., Behrents RG. Sexual dimorphism in normal craniofacial growth. *Angle Orthod*. 1993;63:47-56.

16. Kim YJ, Lee SJ, Kim JW, Jang KT, Lee SH, Hahn SH, Kim CC. Evaluation model of growth and development in pediatric dentistry. *J Korean Acad Pediatr Dent*. 2006;33:173-180.

17. Lee SJ, Kim YJ, Kim TW, Ahn SJ. New evaluation chart of stature and weight for koreans. *Korean J Orthod*. 2006;36:153-160.

18. Park S, Lee JI, Lee SJ, Kim HK, An H. Effect of home environmental characteristics on current growth status and adult stature estimates of sixth-graders. *Korean J Growth Dev*. 2008;16:239-247.

19. Buschang PH, Jacob H, Carrillo R. The morphological characteristics, growth, and etiology of the hyperdivergent phenotype. *Semin Orthod*. 2013;19:212-226.

20. Franchi L, Nieri M, Lomonaco I, McNamara JA, Giuntini V. Predicting the

mandibular growth spurt. *Angle Orthod.* 2021;91:307-312.

21. Kuc-Michalska M, Baccetti T. Duration of the pubertal peak in skeletal class i and class iii subjects. *Angle Orthod.* 2010;80:54-57.

22. Moorrees CF, uan Venrooij ME, Lebret LM, Glatky CG, Kent RL, Reed RB. New norms for the mesh diagram analysis. *Am J Orthod*. 1976;69:57-71.

23. Harris JE, Johnston L, Moyers RE. A cephalometric template: Its construction and clinical significance. *Am J Orthod*. 1963;49:249-263.

24. Ricketts RM. Planning treatment on the basis of the facial pattern and an estimate of its growth. *Angle Orthod.* 1957;27:14-37.

25. Bhatia SN, Wright GW, Leighton BC. A proposed multivariate model for prediction of facial growth. *Am J Orthod*. 1979;75:264-281.

26. Lee HJ, Suh HY, Lee YS, Lee SJ, Donatelli RE, Dolce C, Wheeler TT. A better statistical method of predicting postsurgery soft tissue response in class ii patients. *Angle Orthod.* 2014;84:322-328.

27. Lee YS, Suh HY, Lee SJ, Donatelli RE. A more accurate soft-tissue prediction model for class iii 2-jaw surgeries. *Am J Orthod Dentofacial Orthop*. 2014;146:724-733.

28. Suh HY, Lee HJ, Lee YS, Eo SH, Donatelli RE, Lee SJ. Predicting soft tissue changes after orthognathic surgery: The sparse partial least squares method. *Angle Orthod.* 2019;89:910-916.

29. Suh HY, Lee SJ, Lee YS, Donatelli RE, Wheeler TT, Kim SH, Eo SH, Seo BM. A more accurate method of predicting soft tissue changes after mandibular setback surgery. *J Oral Maxillofac Surg*. 2012;70:e553-562.

30. Yoon KS, Lee HJ, Lee SJ, Donatelli RE. Testing a better method of predicting postsurgery soft tissue response in class ii patients: A prospective study and validity assessment. *Angle Orthod.* 2015;85:597-603.

31. Park JH, Hwang HW, Moon JH, Yu Y, Kim H, Her SB, Srinivasan G, Aljanabi MNA, Donatelli RE, Lee SJ. Automated identification of cephalometric landmarks: Part 1-comparisons between the latest deep-learning methods yolov3 and ssd. *Angle Orthod.* 2019;89:903-909.

32. Hwang HW, Park JH, Moon JH, Yu Y, Kim H, Her SB, Srinivasan G, Aljanabi MNA, Donatelli RE, Lee SJ. Automated identification of cephalometric landmarks: Part 2- might it be better than human? *Angle Orthod.* 2020;90:69-76.

33. Moon JH, Hwang HW, Lee SJ. Evaluation of an automated superimposition method for computer-aided cephalometrics. *Angle Orthod.* 2020;90:390-396.

34. Moon JH, Hwang HW, Yu Y, Kim MG, Donatelli RE, Lee SJ. How much deep learning is enough for automatic identification to be reliable? *Angle Orthod.* 2020;90:823-830.

35. Hwang HW, Moon JH, Kim MG, Donatelli RE, Lee SJ. Evaluation of automated cephalometric analysis based on the latest deep learning method. *Angle Orthod.* 2021;91:329-335.

36. Kim MG, Moon JH, Hwang HW, Cho SJ, Donatelli RE, Lee SJ. Evaluation of an automated superimposition method based on multiple landmarks for growing patients. *Angle Orthod.* 2022;92:226-232.

37. Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. *Imaging Sci Dent.* 2019;49:1-7.

38. Arik SÖ, Pfister T. Tabnet: Attentive interpretable tabular learning. *Proc Conf AAAI Artif Intell*. 2021;35:6679-6687.
39. Sato K, Mito T, Mitani H. An accurate method of predicting mandibular growth potential based on bone maturity. *Am J Orthod Dentofacial Orthop.* 2001;120:286-293.
40. Mito T, Sato K, Mitani H. Predicting mandibular growth potential with cervical vertebral bone age. *Am J Orthod Dentofacial Orthop.* 2003;124:173-177.
41. Chen F, Terada K, Hanada K. A special method of predicting mandibular growth potential for class iii malocclusion. *Angle Orthod*. 2005;75:191-195.
42. Kim JY, Lee SJ, Kim TW, Nahm DS, Chang YI. Classification of the skeletal variation in normal occlusion. *Angle Orthod*. 2005;75:311-319.
43. Lee CH, Park HH, Seo BM, Lee SJ. Modern trends in class iii orthognathic treatment: A time series analysis. *Angle Orthod.* 2017;87:269-278.
44. Donatelli RE, Park JA, Mathews SM, Lee SJ. Time series analysis. *Am J Orthod Dentofacial Orthop*. 2022;161:605-608.
45. Lim HW, Park JH, Park HH, Lee SJ. Time series analysis of patients seeking orthodontic treatment at seoul national university dental hospital over the past decade. *Korean J Orthod.* 2017;47:298-305.
46. Huh YJ, Huh KH, Kim HK, Nam SE, Song HY, Lee JH, Park YS. Constancy of the angle between the frankfort horizontal plane and the sella-nasion line: A nine-year longitudinal study. *Angle Orthod*. 2014;84:286-291.
47. Donatelli RE, Lee SJ. How to test validity in orthodontic research: A mixed dentition analysis example. *Am J Orthod Dentofacial Orthop*. 2015;147:272-279.

48. Moon JH, Kim MG, Hwang HW, Cho SJ, Donatelli RE, Lee SJ. Evaluation of an individualized facial growth prediction model based on the multivariate partial least squares method. *Angle Orthod*. 2022;92:705-713.
49. Donatelli RE, Lee SJ. How to report reliability in orthodontic research: Part 1. *Am J Orthod Dentofacial Orthop*. 2013;144:156-161.
50. Donatelli RE, Lee SJ. How to report reliability in orthodontic research: Part 2. *Am J Orthod Dentofacial Orthop*. 2013;144:315-318.
51. Yang L, Shami A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*. 2020;415:295-316.
52. Fernández A, Garcia S, Herrera F, Chawla NV. Smote for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *J Artif Intell Res*. 2018;61:863-905.
53. Team RDC. 2019. R: A language and environment for statistical computing.

Table 1. Summary of Growth Prediction Methods

Research Group	Year	No. of Study Subjects	No. of Prediction Landmarks	Growth Observation Range (years)	Growth Prediction Methods
Present study	2022	303	80	13.2	Multivariate partial least squares regression model; TabNet Deep Neural Network
Moon *et al.*[48]	2022	303	80	13.2	Multivariate partial least squares regression model
Chvatal *et al*.[2]	2005	287	4	9.0	Multilevel model
Chen *et al*.[41]	2005	44	Md length (Ar-Pog)	10.0	Multiple linear regression
Mito *et al*.[40]	2003	40	Md length (Cd-Gn)	NA	Multiple linear regression
Sato *et al*.[39]	2001	44	Md length (Cd-Gn)	10.1	Multiple linear regression
Rudolph *et al*.[13]	1998	31	26	12.0	Multivariate statistical method
Suzuki and Takahama[14]	1991	250	67	6.4	Principal component analysis; Cluster analysis; Discriminant analysis; Multiple linear regression
Bhatia *et al*.[25]	1979	80	12	8.0	Factor analysis; Cluster analysis; Discriminant analysis
Popovich and Thompson[10]	1977	210	20	16.0	Craniofacial templates based on population norm
Moorrees *et al.*[22] Moorrees and Lebret[9]	1976 1962	93	34	NA	Mesh diagram based on population norm
Johnston[5] Harris *et al.*[23]	1975 1963	45	7	12.0	Forecast grid based on population norm
Ricketts[24]	1957	NA	20	NA	Visual treatment objective based on population norm

Md, mandibular; Ar, Articulare; Pog, Pogonion; Cd, Condylion; Gn, Gnathion.

Table 2. Characteristics of Subjects (n = 303)

Variable	N	(%)	Mean	SD	Min	Max
Age (years)						
Female, beginning of growth observation	166	(54.8%)	11.0	3.0	4.2	18.6
Male, beginning of growth observation	137	(45.2%)	10.9	3.0	5.9	18.7
All subjects, beginning of growth observation			10.9	3.0	4.2	18.7
All subjects, end of growth observation			14.2	3.9	6.7	25.6
Growth observation period (years)			3.3	2.6	1.0	13.2
Number of serial radiographs taken						
Two	251	(82.8%)				
Three	39	(12.9%)				
Four	13	(4.3%)				
Molar relationship at first visit						
Class I	78	(25.7%)				
Class II	118	(38.9%)				
Class III	107	(35.3%)				

SD, standard deviation; Min, minimum; Max, maximum.

Table 3. Multiple Linear Regression Analysis of Factors Influencing Growth Prediction Error

Factor	β	SE (β)	*P*-value
Age at prediction (years)	-0.01	0.004	0.0112
Growth prediction interval (years)	0.14	0.004	< 0.0001
Prediction method			
The partial least squares method	*Reference*		
Artificial Intelligence TabNet DNN algorithm	-1.11	0.021	< 0.0001
Sex			
Female	*Reference*		
Male	0.07	0.021	0.0005
Molar relationship			
Class I	*Reference*		
Class II	-0.18	0.028	< 0.0001
Class III	-0.01	0.029	0.7735
Direction of growth			
Anteroposterior direction (*x* axis)	*Reference*		
Vertical direction (*y* axis)	-0.03	0.021	0.1680
Type of landmark			
Hard tissue	*Reference*		
Soft tissue	0.18	0.021	< 0.0001
Position of landmark			
Mandible	*Reference*		
Maxilla	-0.67	0.021	< 0.0001

β, regression coefficients; SE, standard error; DNN, deep neural network.

Table 4. Comparison of growth prediction models based on the partial least squares (PLS) method and the TabNet artificial intelligence (AI) algorithm. Values are the Euclidean distance between prediction results and actual growth in millimeter unit. For a given landmark, the model that showed more accurate prediction results is indicated by symbol √

	PLS Method				TabNet AI algorithm				More Accurate		
Landmark[a]	Mean	SD	Min	Max	Mean	SD	Min	Max	PLS	AI	*P* value[b]
Nasion	1.0	0.8	0.0	4.9	1.7	1.3	0.0	6.8	√		<0.0001
Nasal tip	2.5	1.4	0.2	12.3	2.7	1.5	0.1	10.9			1.0000
Porion	1.9	1.2	0.0	13.0	3.0	1.6	0.1	12.1	√		<0.0001
Orbitale	2.0	1.2	0.1	11.7	2.4	1.3	0.1	7.1	√		0.0001
Anterior nasal spine	2.6	1.8	0.1	22.0	2.1	1.3	0.1	7.1		√	<0.0001
Posterior nasal spine	2.4	1.6	0.1	18.5	2.3	1.3	0.1	8.0			1.0000
Point A	2.8	2.0	0.1	28.3	2.1	1.2	0.1	8.9		√	<0.0001
U1 root tip	2.9	2.1	0.0	32.1	2.1	1.2	0.1	6.7		√	<0.0001
U1 incisal edge	3.8	2.6	0.1	37.7	2.8	1.6	0.1	9.3		√	<0.0001
L1 incisal edge	3.9	2.9	0.2	41.7	2.5	1.5	0.0	9.3		√	<0.0001
L1 root tip	4.4	3.5	0.3	49.5	2.2	1.4	0.1	10.2		√	<0.0001
Point B	4.6	3.7	0.2	52.5	2.2	1.3	0.2	9.7		√	<0.0001
Protuberance menti	4.8	3.9	0.1	55.5	2.1	1.4	0.2	13.0		√	<0.0001
Pogonion	5.2	4.2	0.1	58.7	2.4	1.6	0.1	14.0		√	<0.0001
Gnathion	5.2	4.3	0.2	59.4	2.4	1.5	0.1	13.0		√	<0.0001
Menton	5.2	4.3	0.2	60.1	2.4	1.5	0.0	11.2		√	<0.0001
Gonion, constructed	4.0	3.0	0.2	37.3	3.1	1.9	0.1	13.1		√	<0.0001
Gonion, anatomic	4.0	3.0	0.1	37.1	3.1	1.9	0.1	10.8		√	<0.0001
Articulare	2.0	1.3	0.1	14.3	2.8	1.6	0.1	8.8	√		<0.0001
Condylion	1.7	1.1	0.1	12.6	2.8	1.6	0.3	11.2	√		<0.0001
Pterygoid	1.8	1.1	0.1	10.8	2.5	1.4	0.1	8.0	√		<0.0001
Basion	2.5	1.6	0.1	16.5	3.3	1.9	0.1	11.0	√		<0.0001
glabella	3.4	2.3	0.1	13.8	4.2	2.6	0.1	20.3	√		<0.0001
nasion	2.2	1.2	0.0	7.4	2.6	1.4	0.1	8.9	√		<0.0001
supranasal tip	3.0	1.8	0.2	16.0	2.6	1.4	0.1	10.9		√	0.0263
pronasale	3.1	1.9	0.1	19.7	2.4	1.3	0.0	8.5		√	<0.0001
columella	3.1	2.0	0.2	23.4	2.3	1.3	0.1	7.6		√	<0.0001
subnasale	3.0	1.9	0.1	24.0	2.0	1.2	0.2	7.4		√	<0.0001
point A	3.0	1.9	0.0	25.0	1.9	1.1	0.1	6.5		√	<0.0001
superior labial sulcus	3.3	2.3	0.0	31.1	2.1	1.2	0.1	7.1		√	<0.0001
labiale superius	3.5	2.4	0.1	31.9	2.4	1.4	0.1	8.6		√	<0.0001
upper lip	3.6	2.6	0.3	33.1	2.4	1.4	0.1	8.7		√	<0.0001
stomion superius	3.8	2.7	0.0	35.5	2.3	1.4	0.1	10.4		√	<0.0001
stomion inferius	4.1	2.9	0.3	38.5	2.6	1.5	0.1	11.4		√	<0.0001
lower lip	4.6	3.2	0.2	41.8	2.7	1.6	0.1	10.3		√	<0.0001
labiale inferius	4.8	3.3	0.2	41.9	2.9	1.8	0.2	12.8		√	<0.0001
point B	4.9	3.6	0.1	50.3	2.8	1.6	0.3	9.9		√	<0.0001
protuberance menti	5.2	3.7	0.2	50.0	2.7	1.7	0.1	10.8		√	<0.0001
pogonion	5.5	3.9	0.1	47.6	3.1	2.0	0.1	13.0		√	<0.0001
gnathion	5.5	4.4	0.4	59.2	2.9	1.8	0.2	12.6		√	<0.0001
menton	5.5	4.6	0.2	62.5	2.6	1.6	0.2	11.3		√	<0.0001

[a] In this table, landmarks were selected to show the results succinctly. Soft tissue landmarks were indicated by small case letters, and hard tissue landmarks by capital letters.

[b] Results from *t*-tests with Bonferroni correction. SD, standard deviation; Min, minimum; Max, maximum.

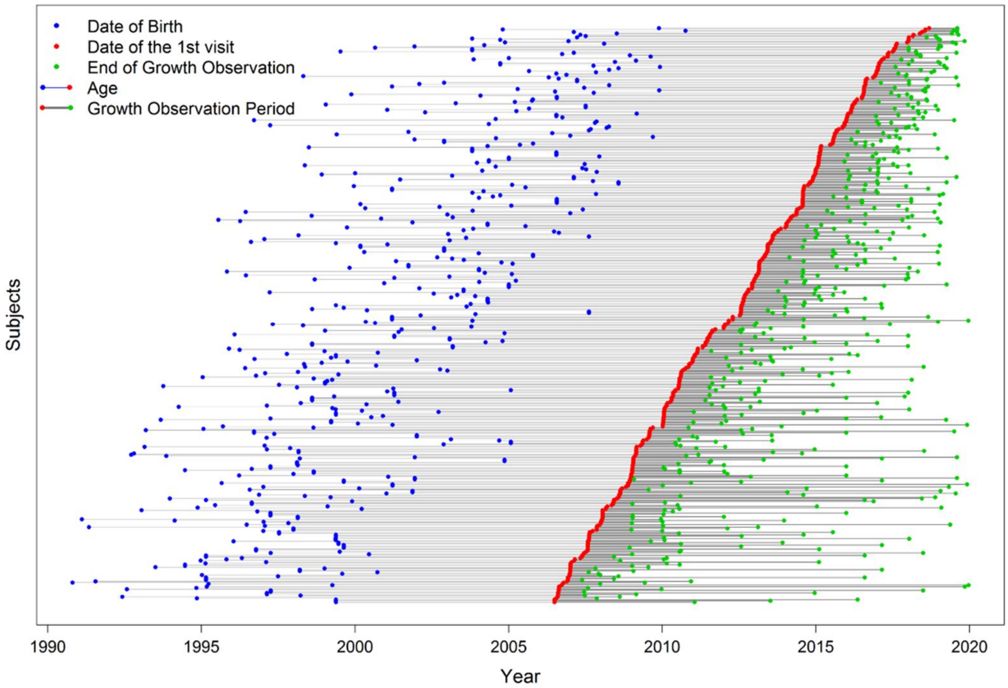


Figure 1. Growth observation period for each subject. Red and green dots indicate when the first and second cephalometric images were taken, respectively.

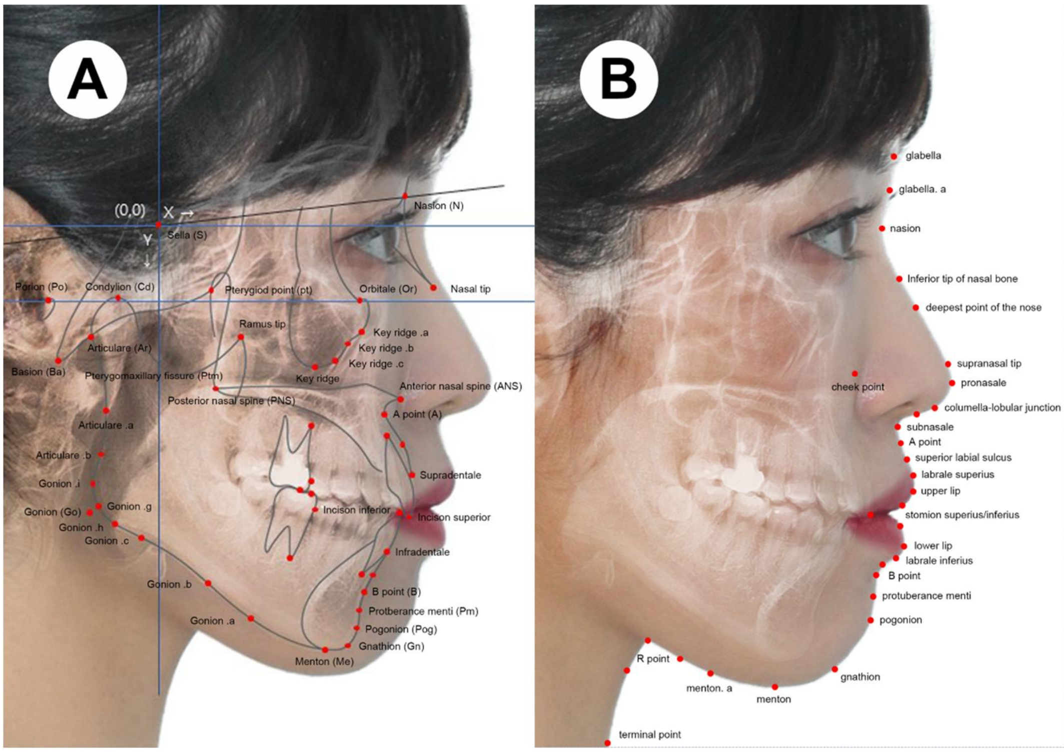


Figure 2. Reference planes and cephalometric landmarks used in present study. A, Skeletal landmarks are shown in capital letters. B, Soft tissue landmarks are presented in lowercase letters.

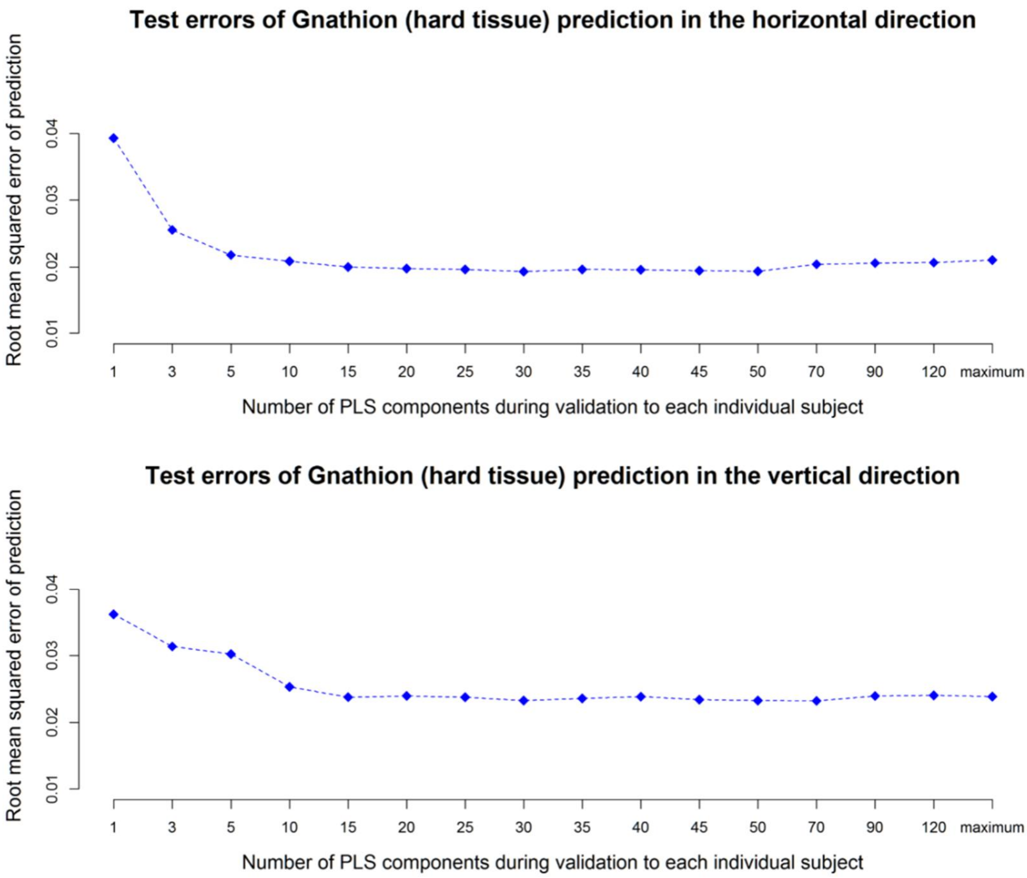


Figure 3. Growth prediction error according to the number of PLS components. Growth prediction errors for Gnathion were chosen to show the pattern of error in the horizontal direction (top) and the vertical direction (bottom).

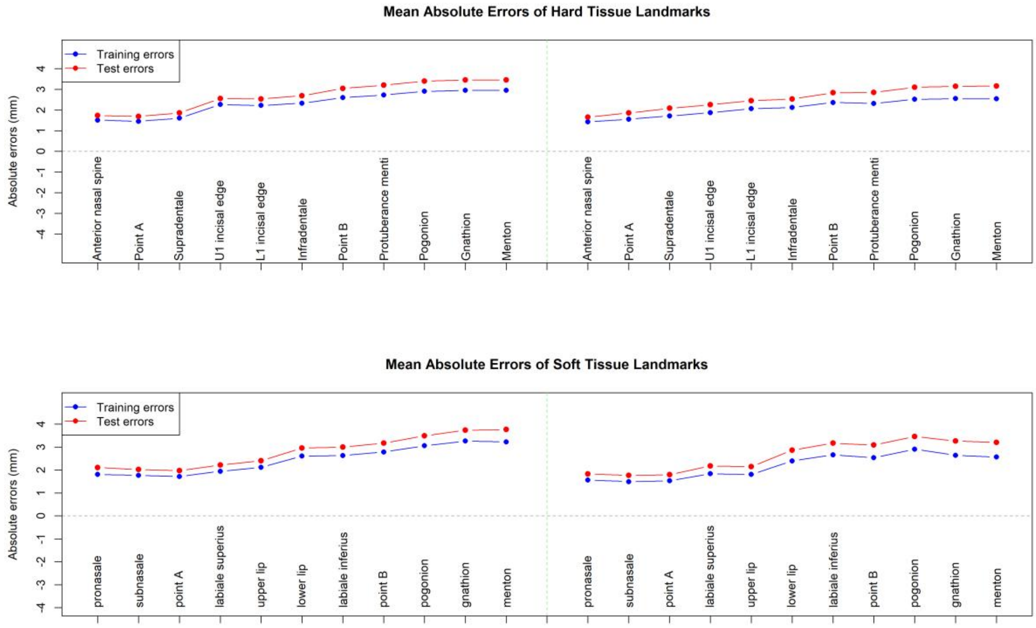


Figure 4. Growth prediction errors for selected landmarks in the training data set (*blue*) and the test data set (*red*).

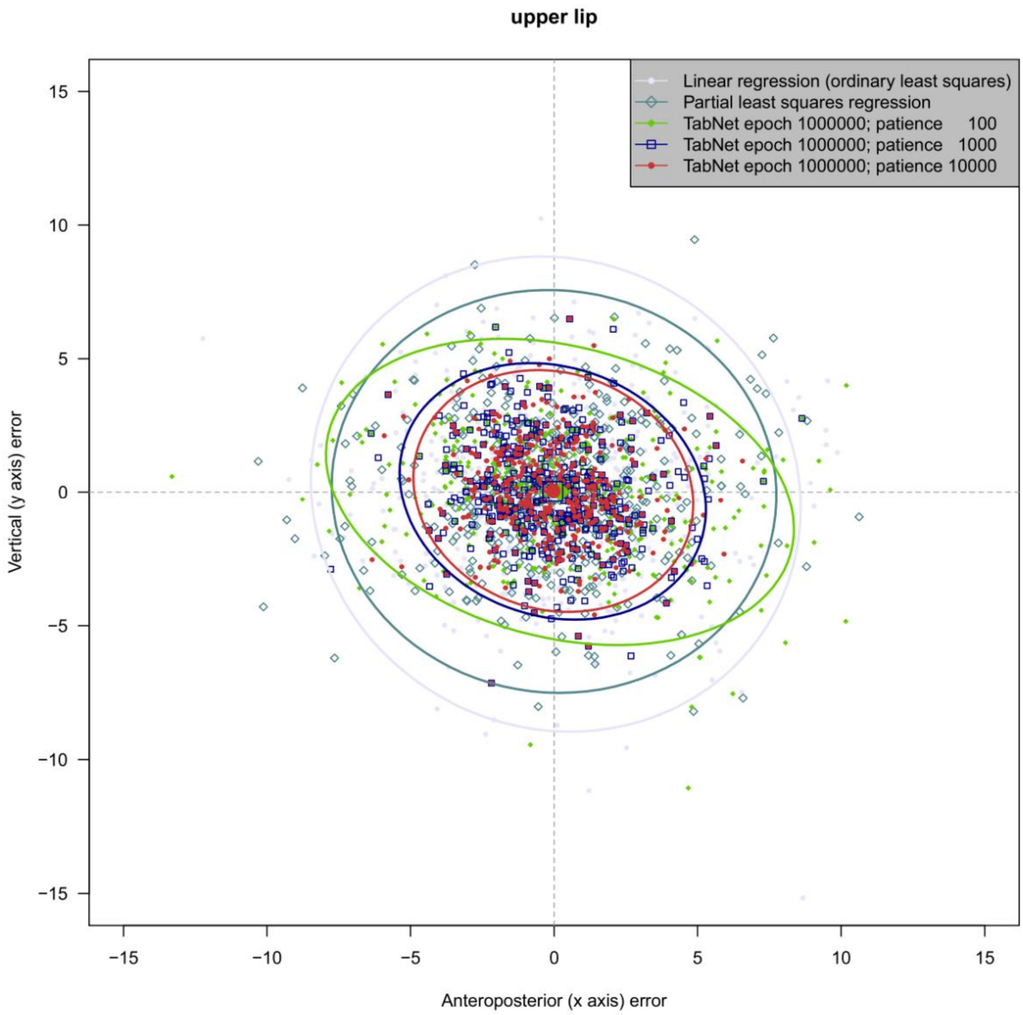


Figure 5. The process of hyperparameter tuning. Growth prediction errors of AI-based prediction models were graphically compared to find optimal hyperparameter values. Prediction errors for “upper lip” were chosen as an example.

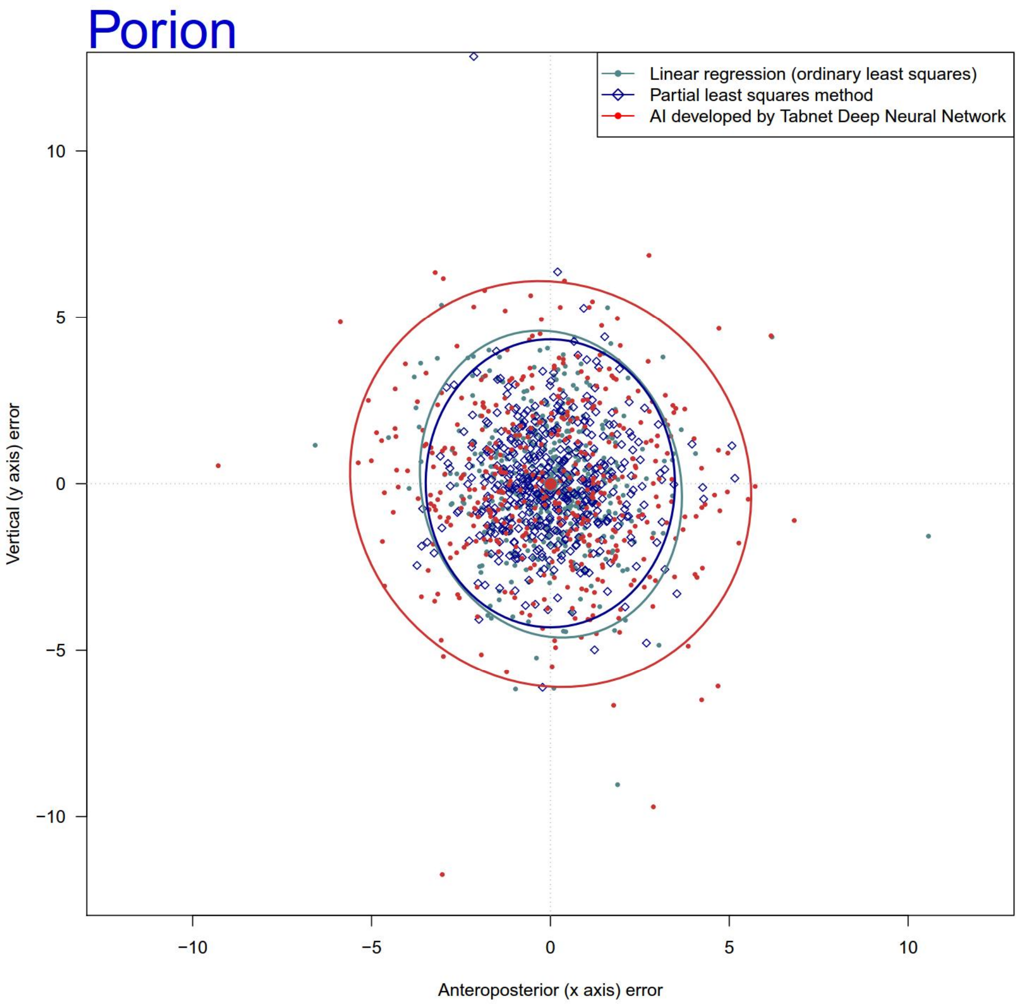


Figure 6, A. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Porion" case.

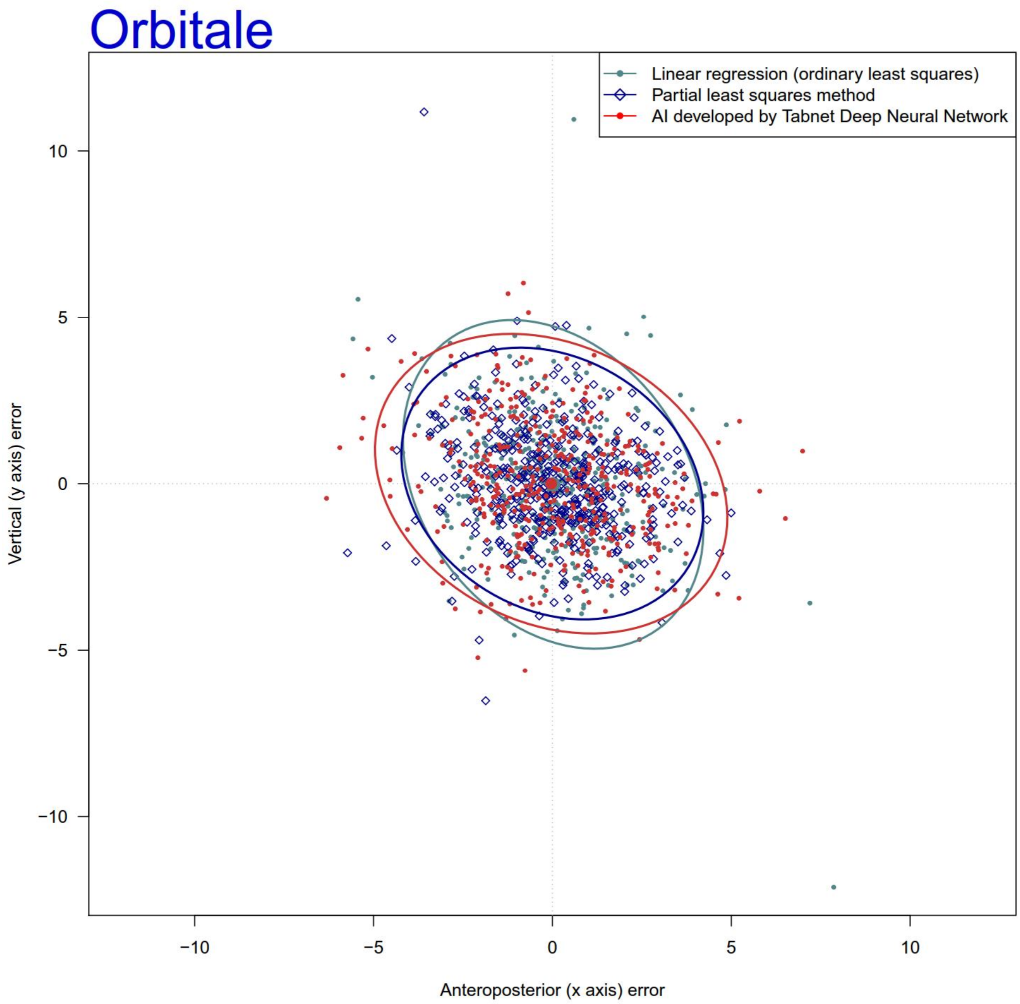


Figure 6, B. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Orbitale" case.

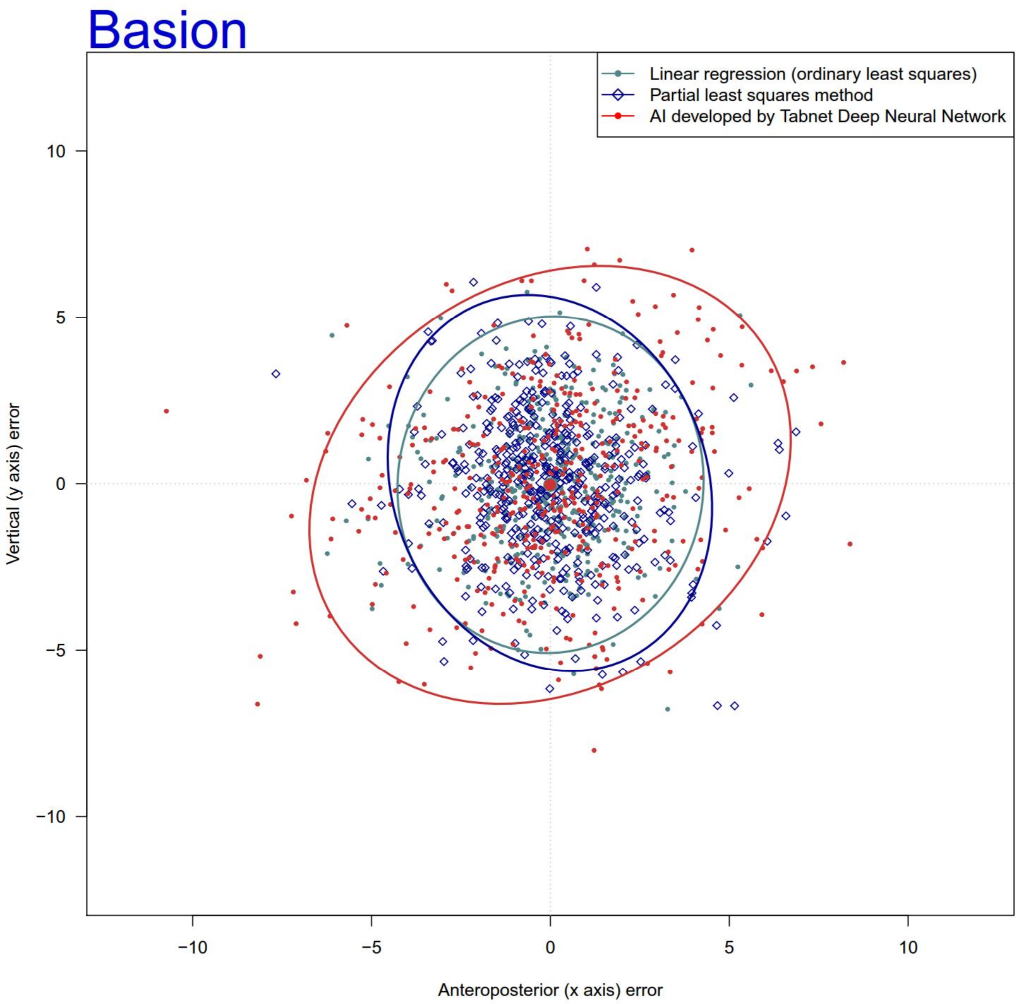


Figure 6, C. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Basion" case.

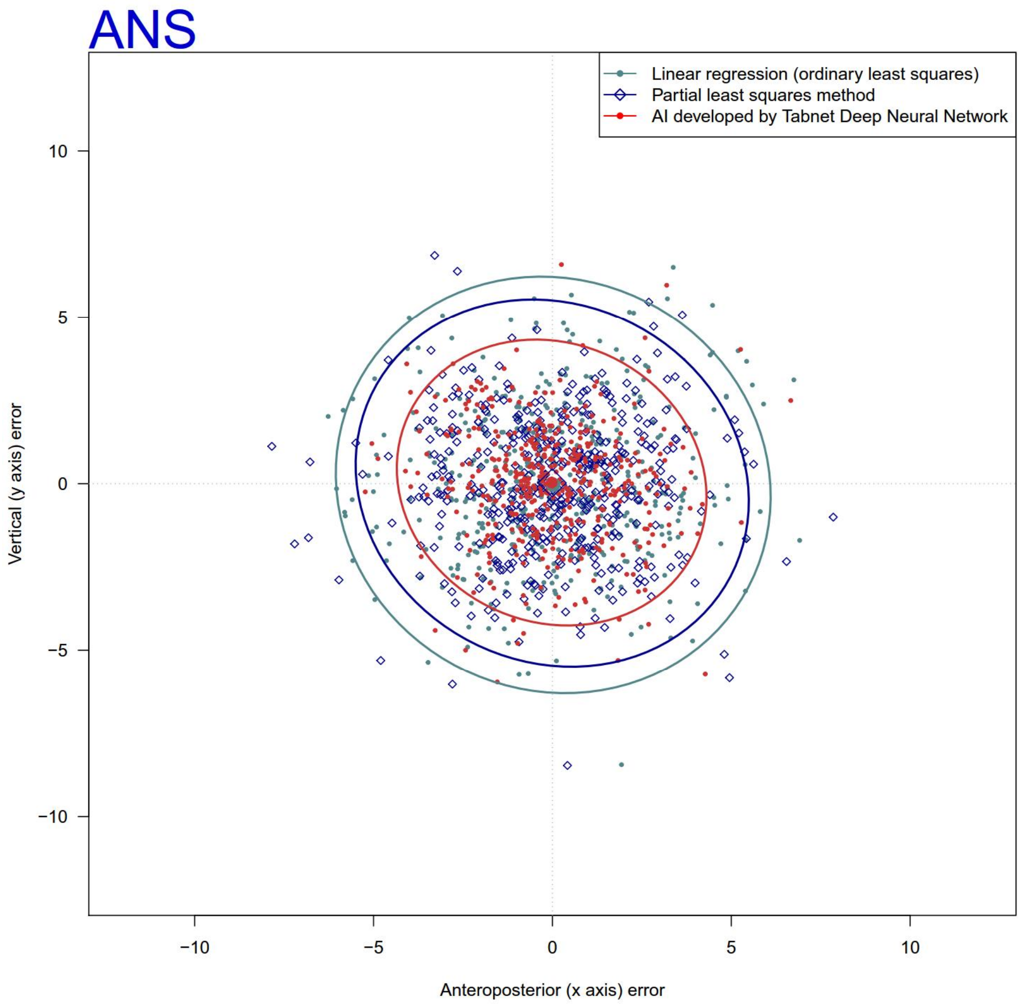


Figure 6, D. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Anterior nasal spine" case.

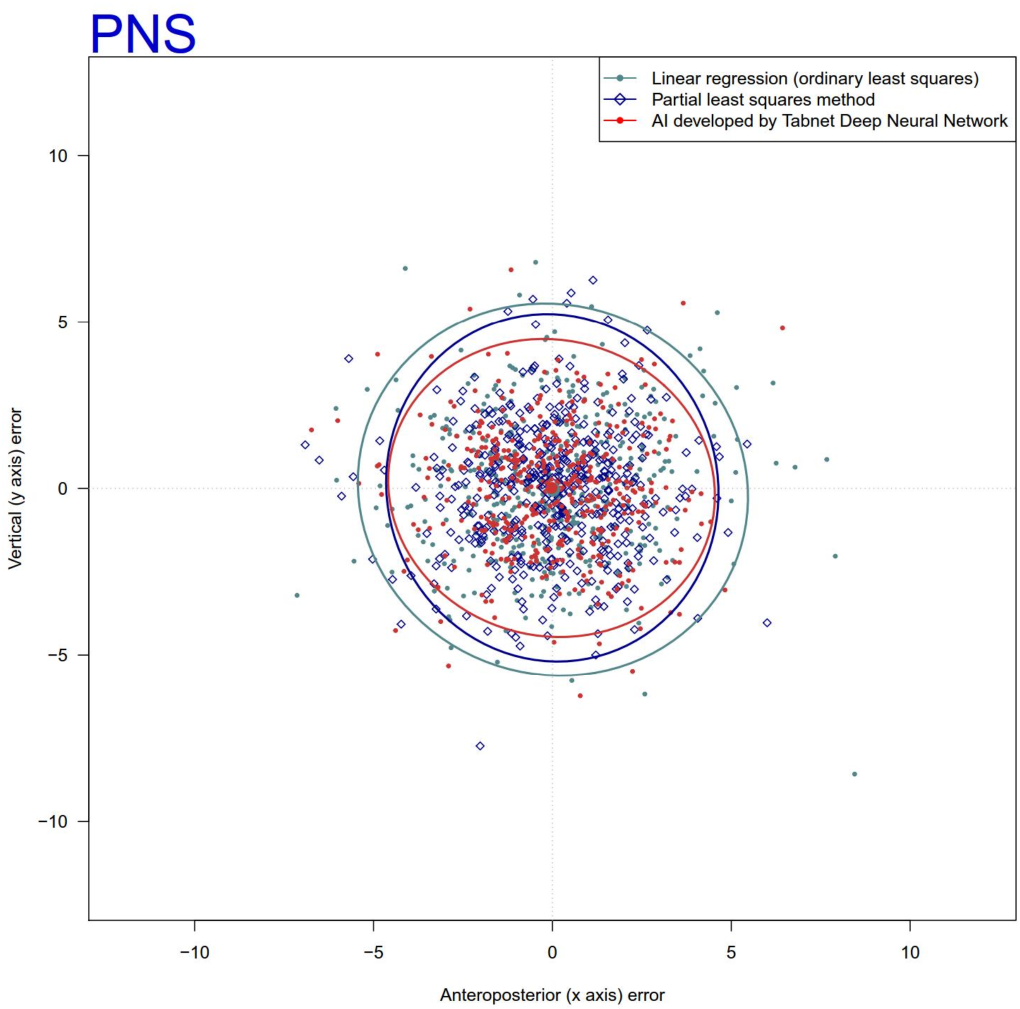


Figure 6, E. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the " Posterior nasal spine " case.

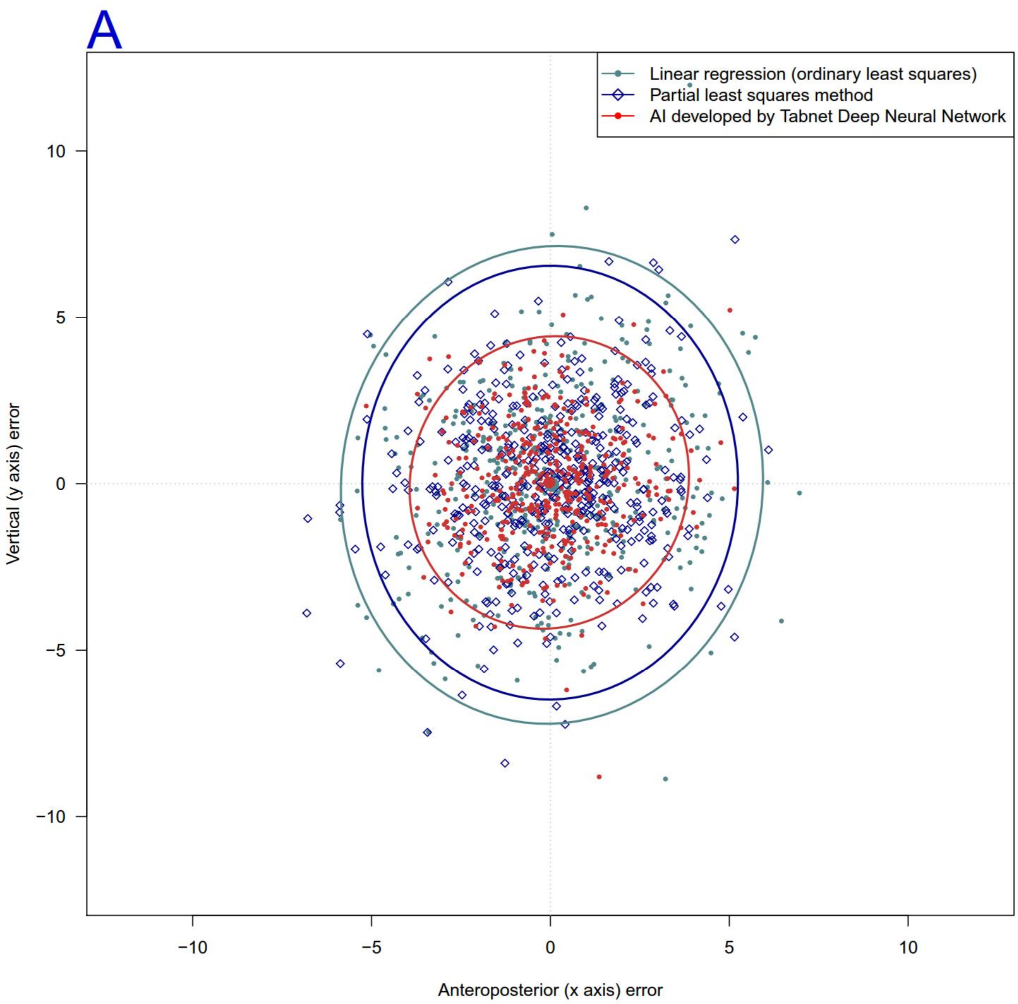


Figure 6, F. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Point A" case.

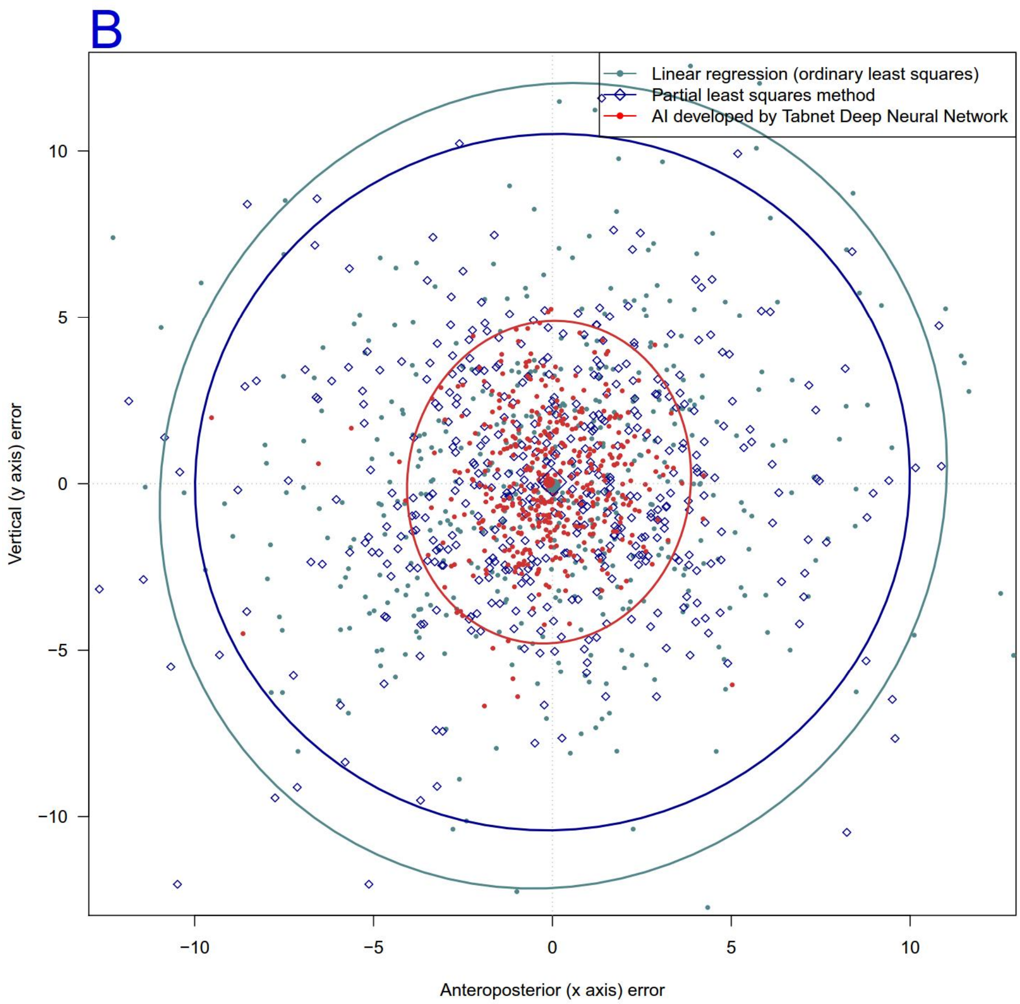


Figure 6, G. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Point B" case.

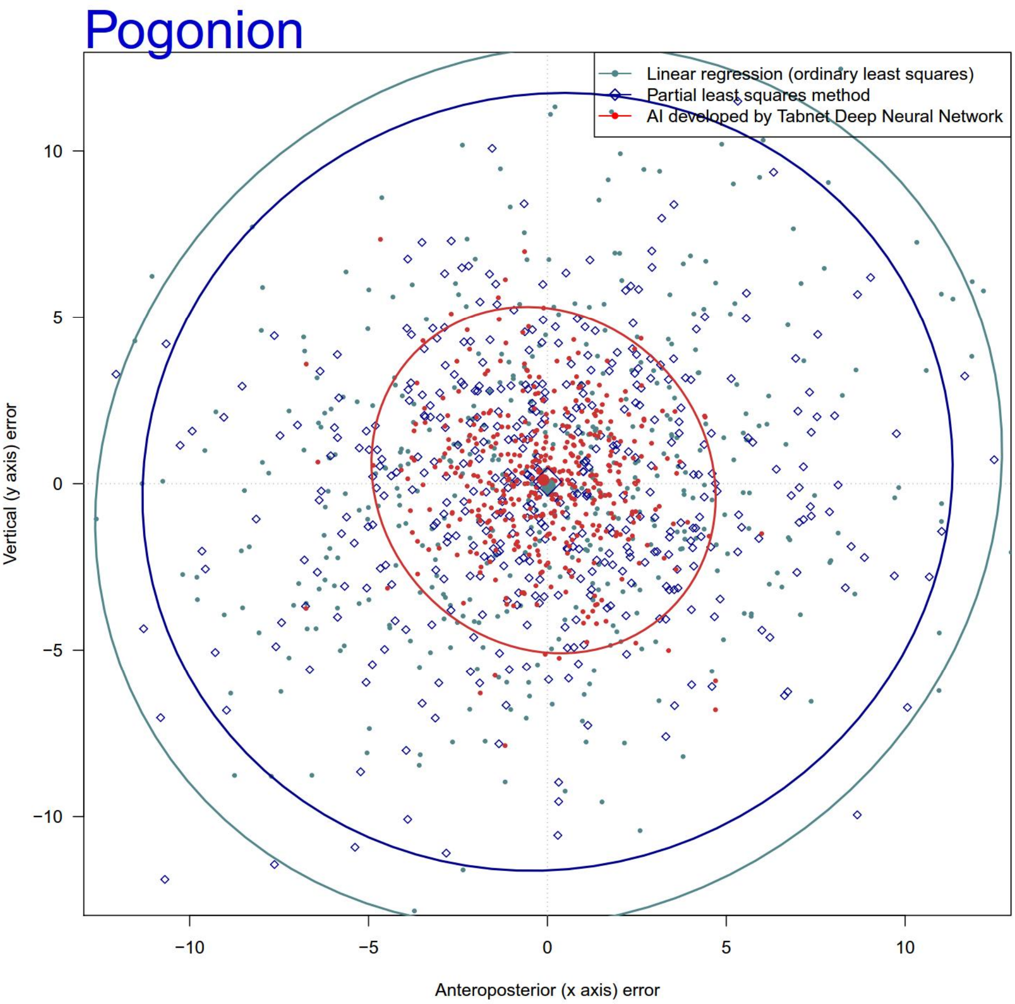


Figure 6, H. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Pogonion" case.

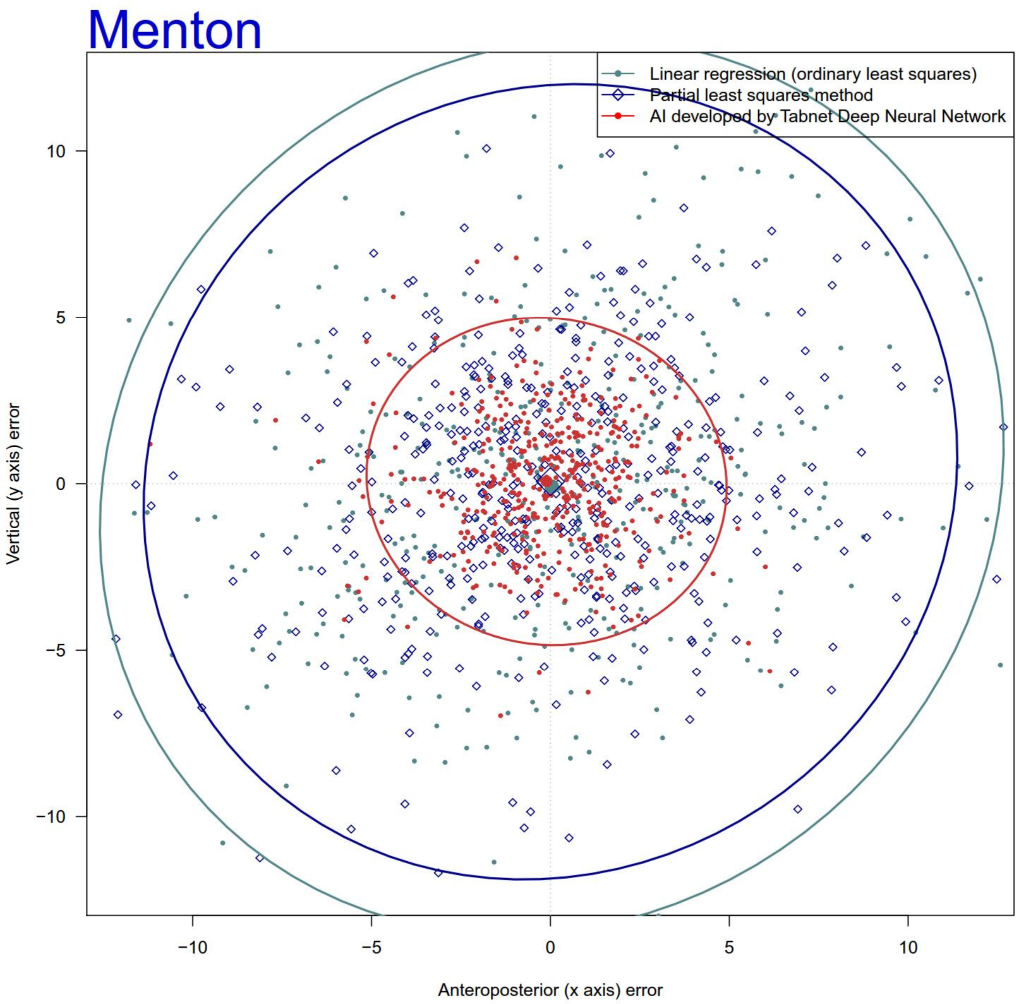


Figure 6, I. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "Menton" case.

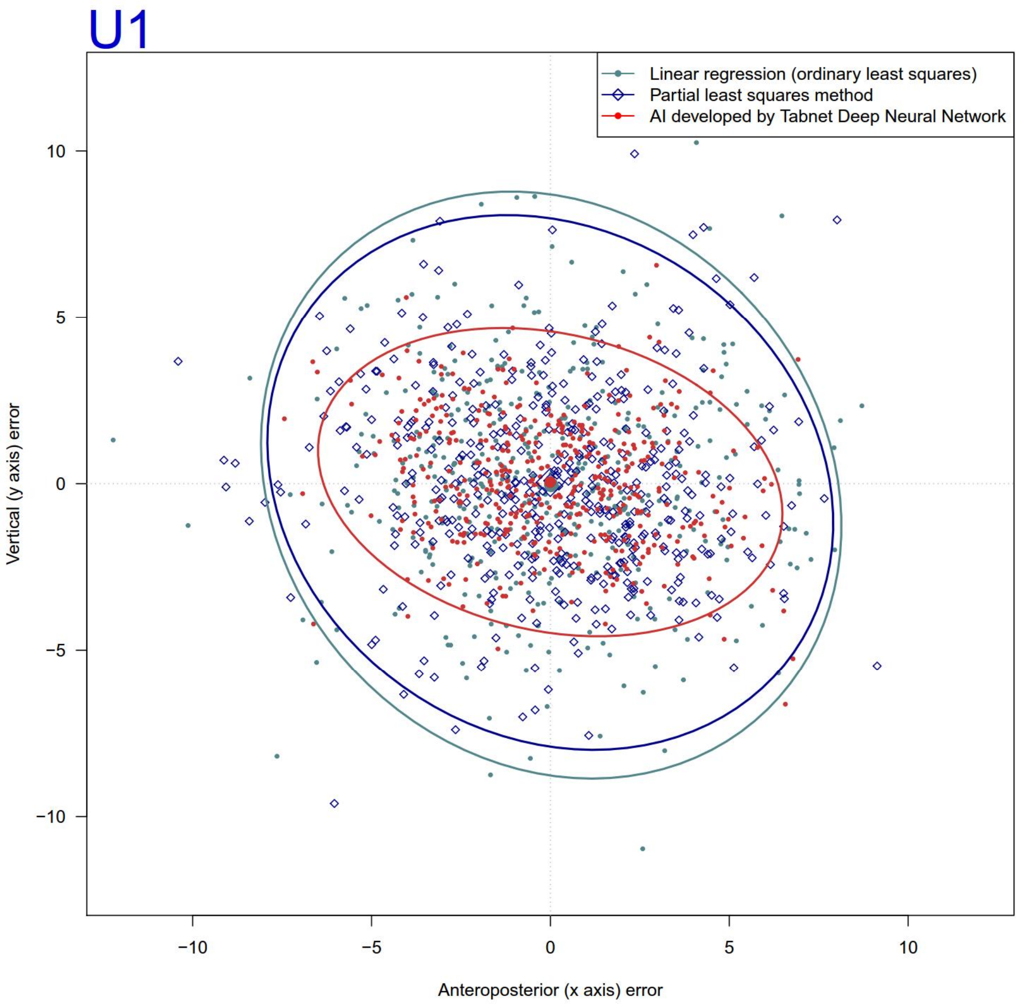


Figure 6, J. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "U1 incisal edge" case.

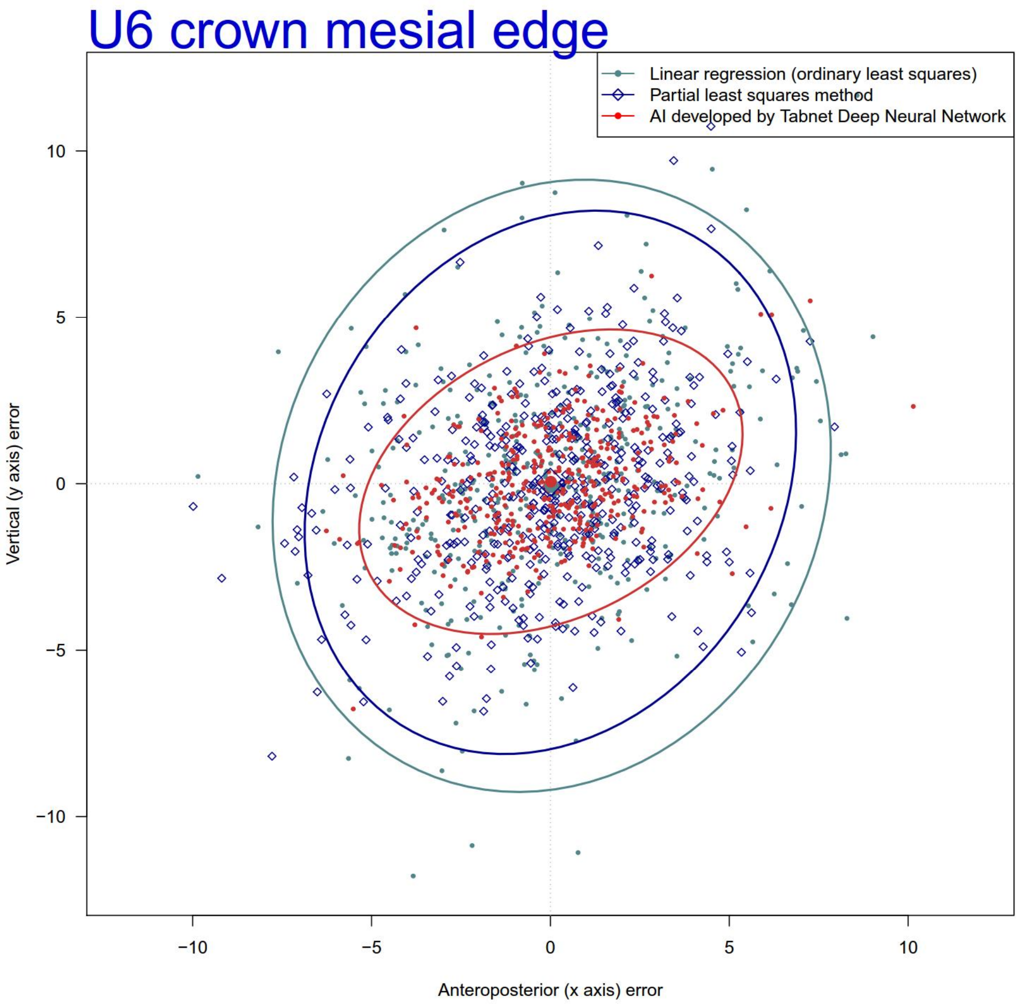


Figure 6, K. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "U6 mesial contact point" case.

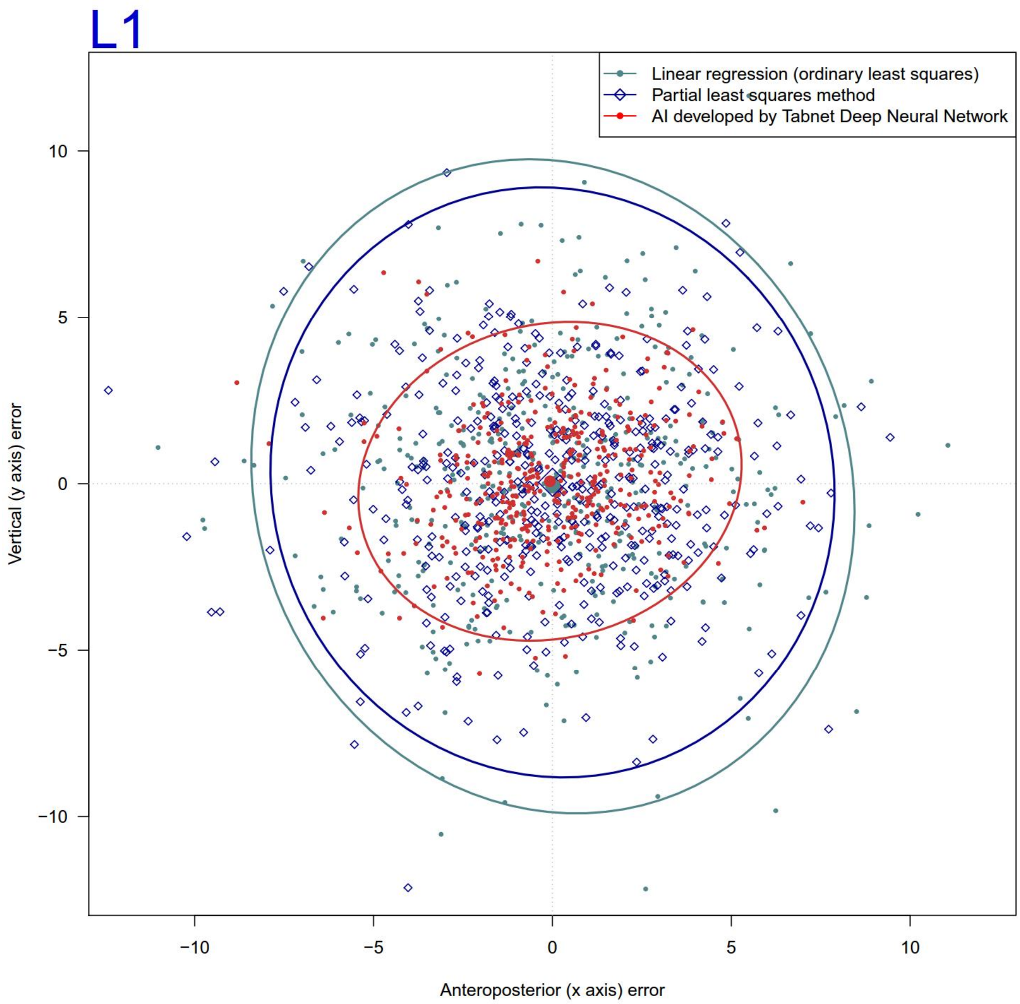


Figure 6, L. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "L1 incisal edge" case.

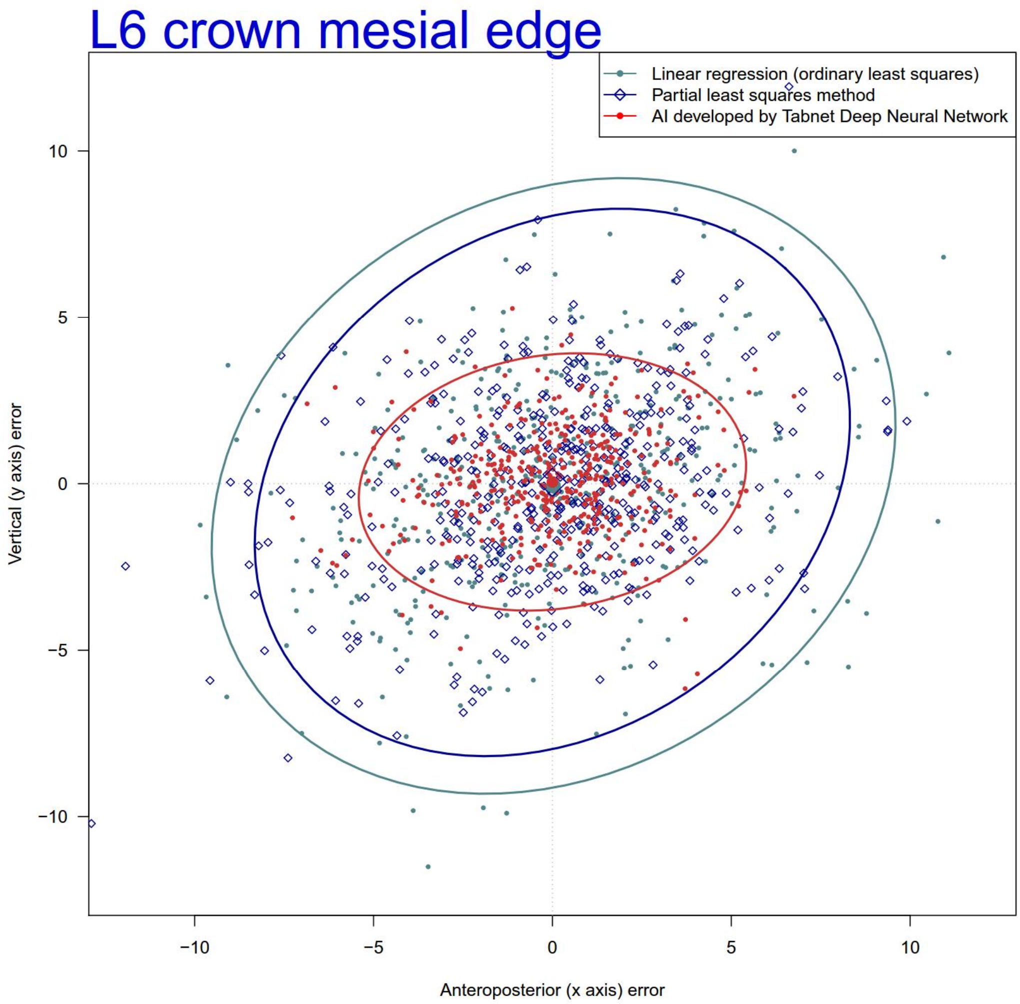


Figure 6, M. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "L6 mesial contact point" case.

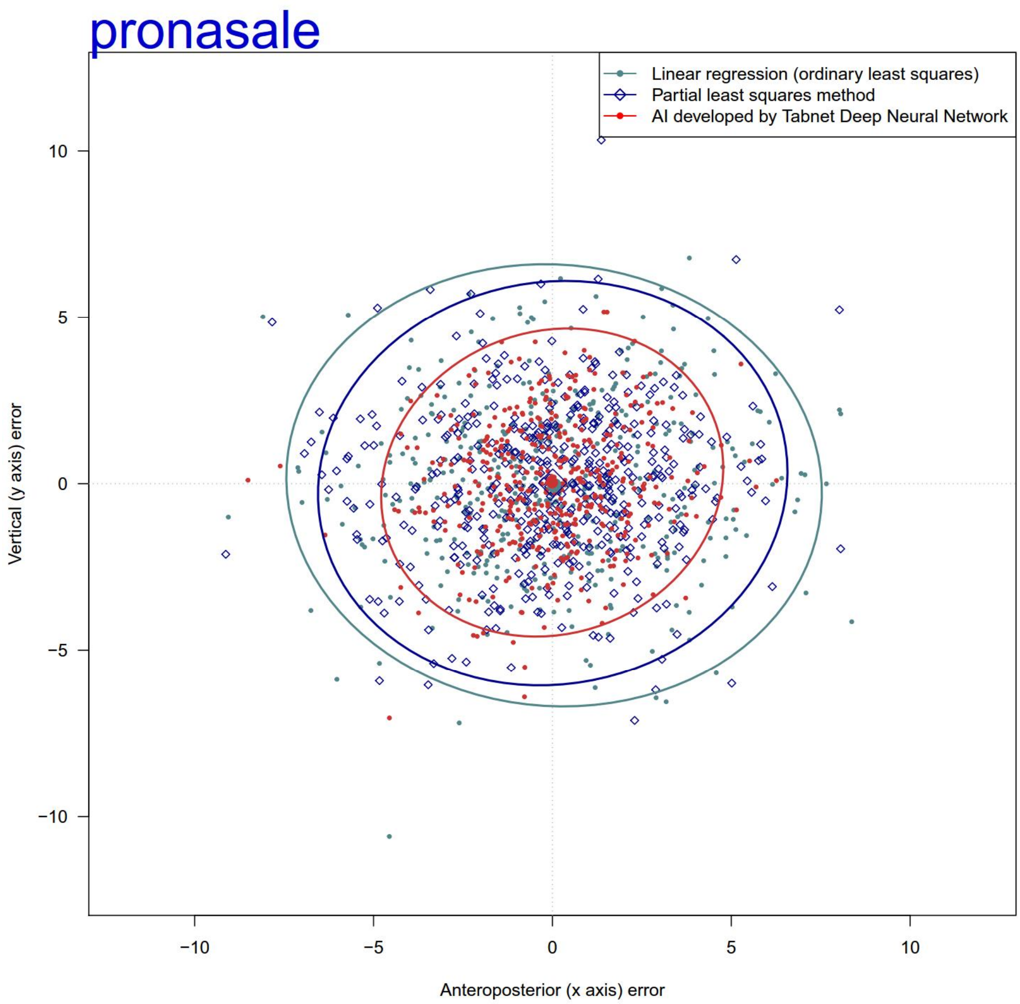


Figure 6, N. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "pronasale" case.

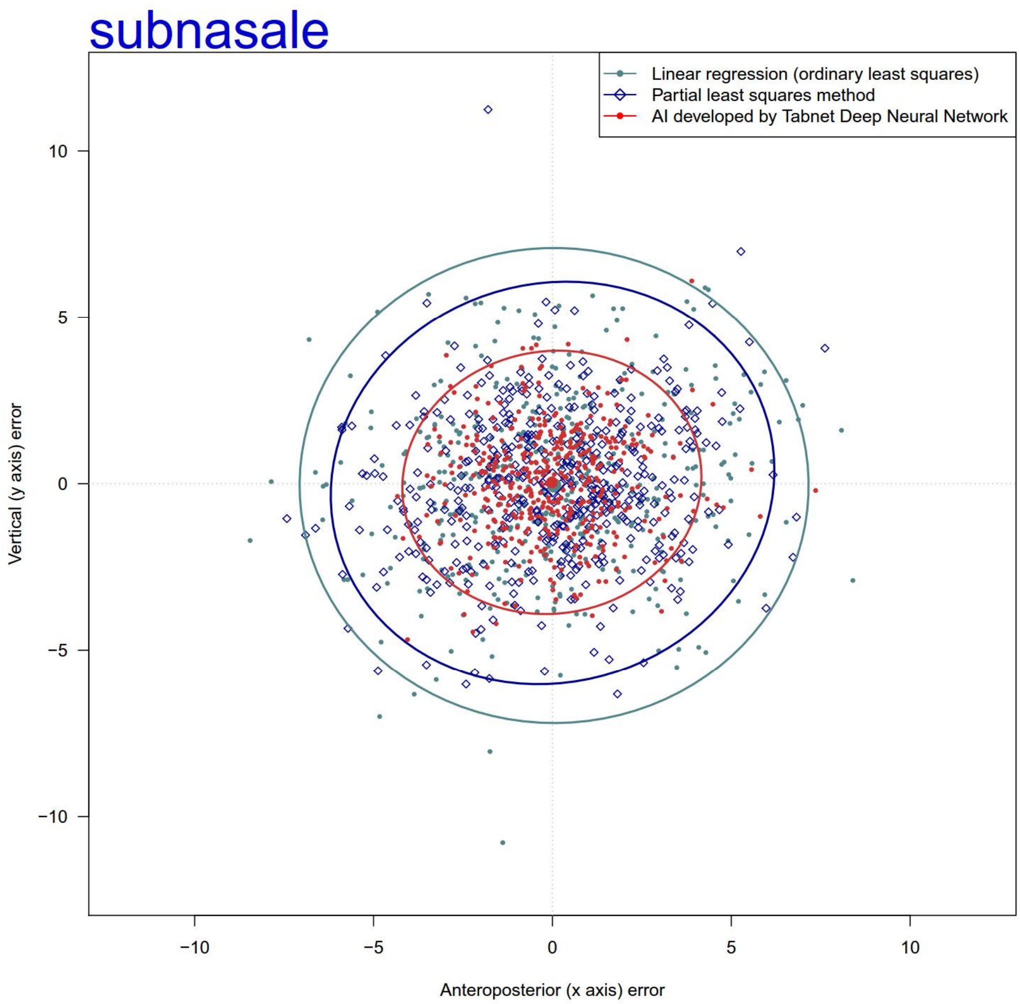


Figure 6, O. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "subnasale" case.

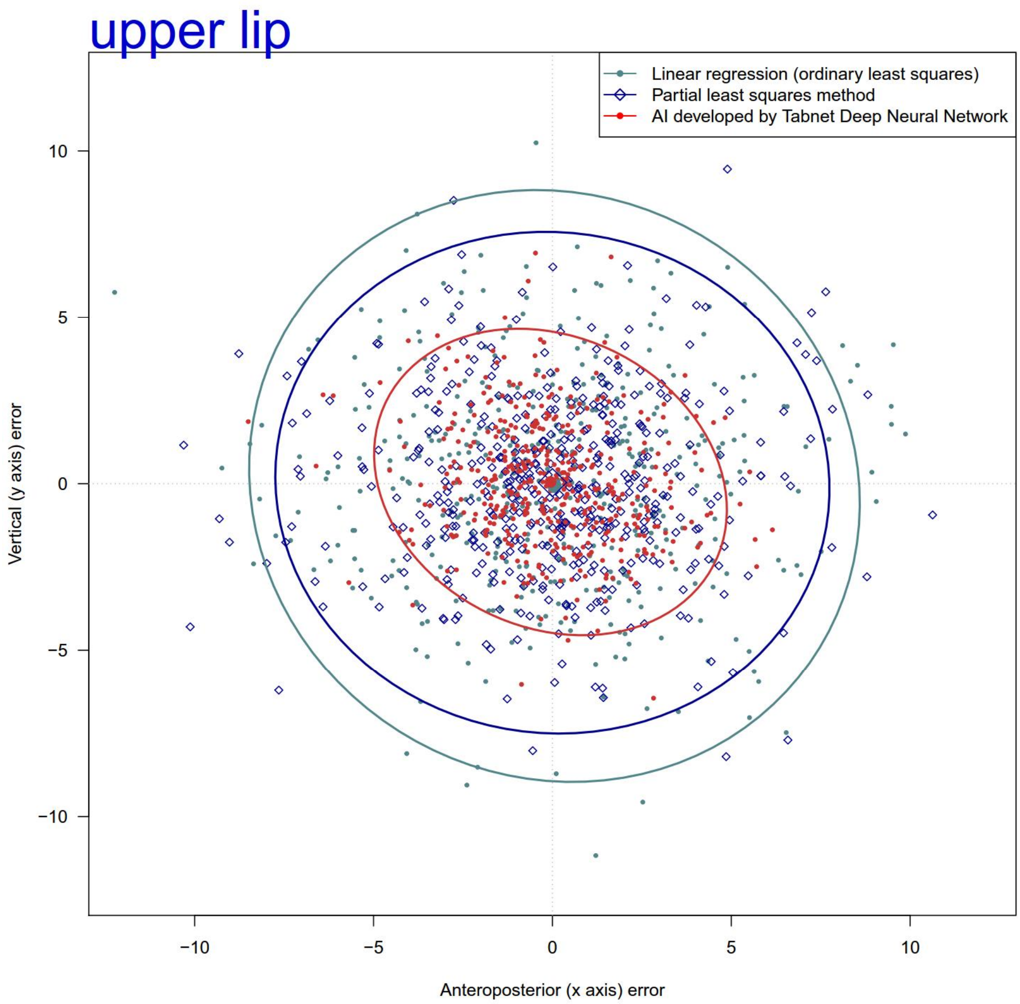


Figure 6, P. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "upper lip" case.

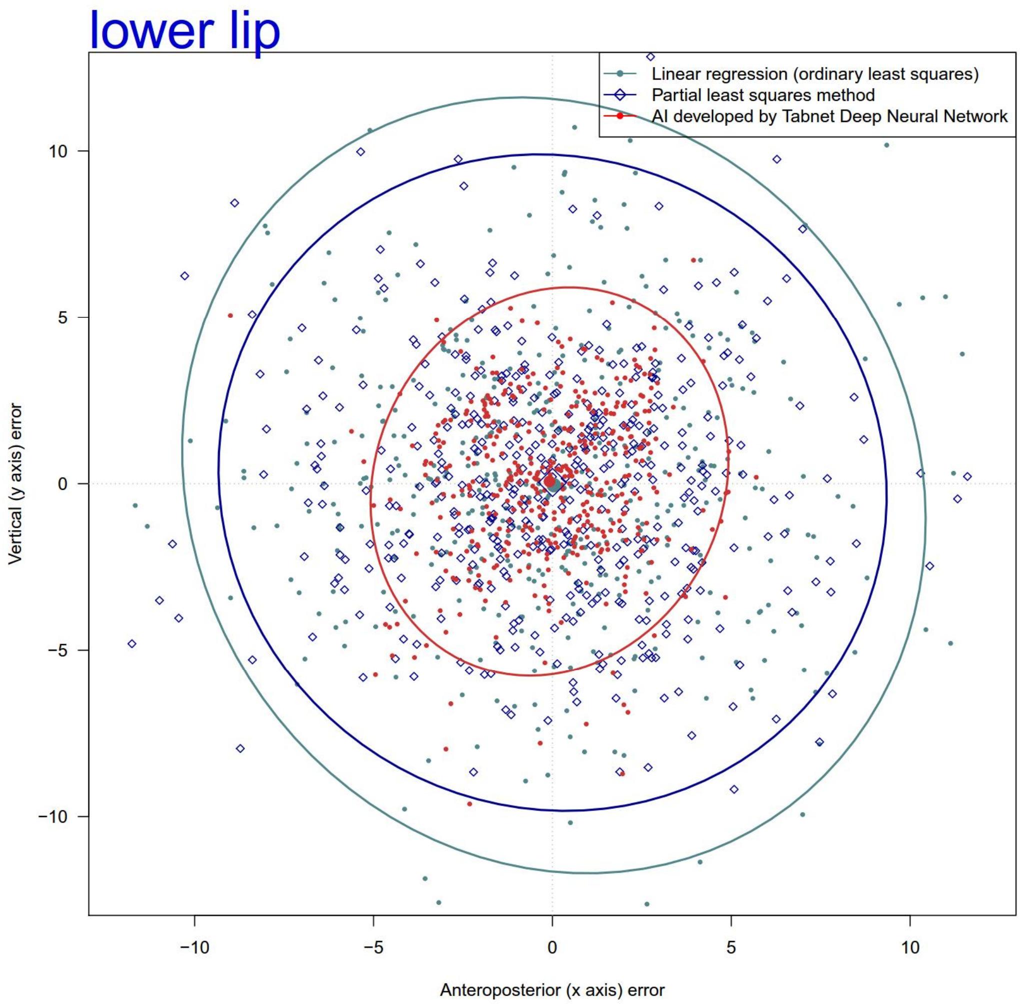


Figure 6, Q. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the " lower lip" case.

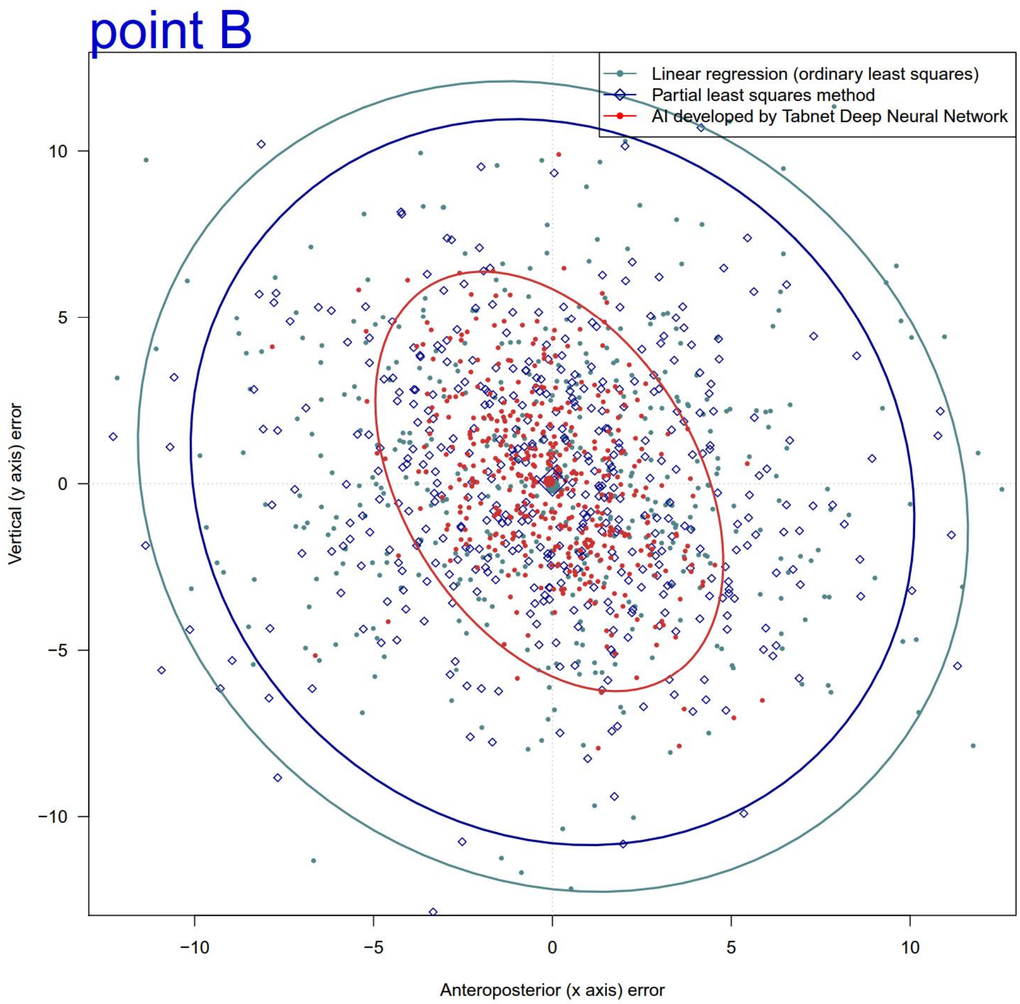


Figure 6, R. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "point B" case.

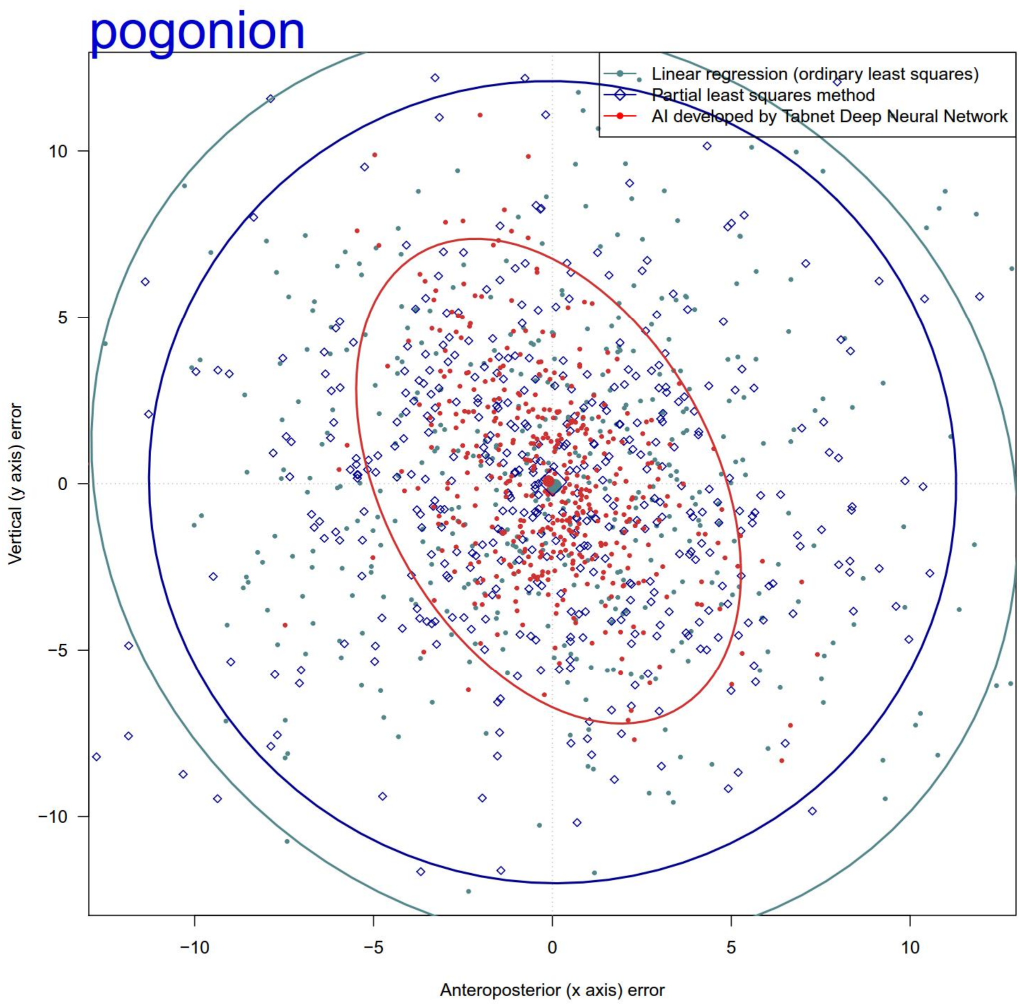
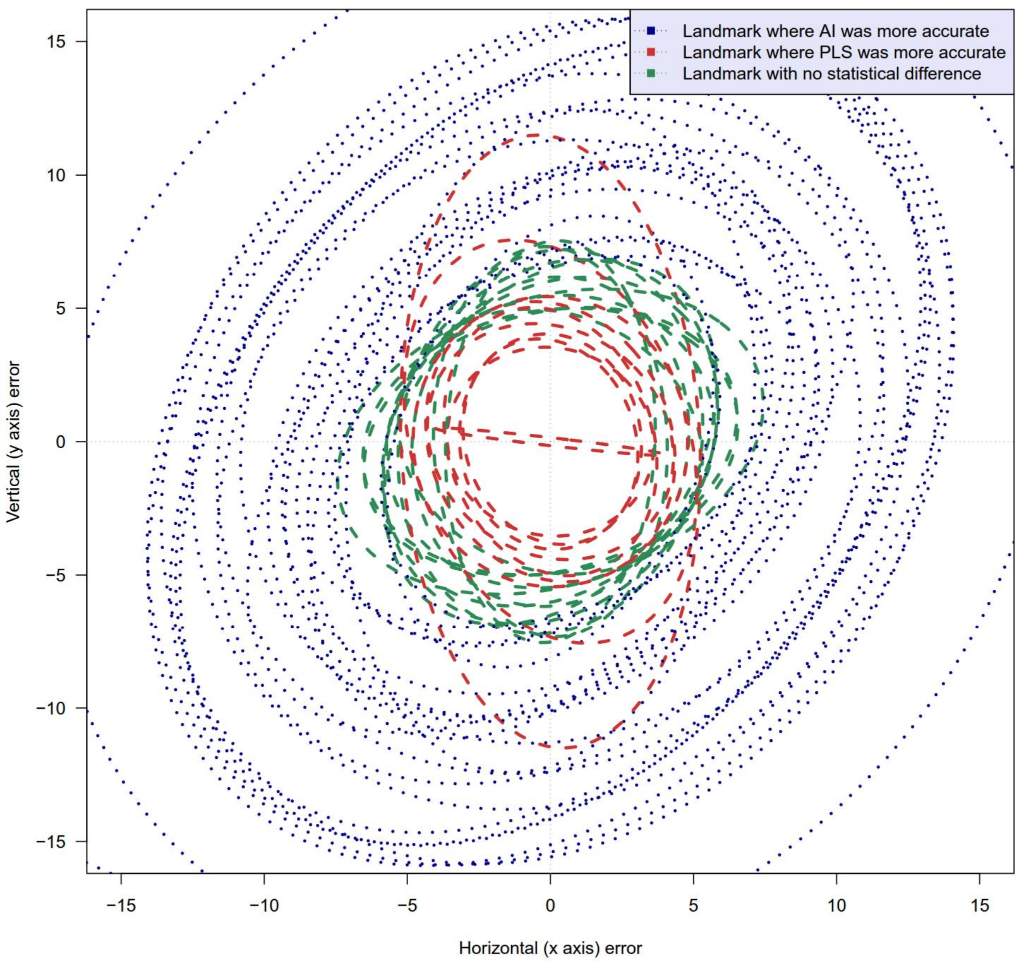


Figure 6, S. Scatterplots and 95% confidence ellipse illustrating the pattern of growth prediction errors in the "pogonion" case.

Figure 7. Growth pattern and variability of each landmark according to the model that showed more accurate predictions. Landmarks where PLS-based predictions were superior tended to have less variability in growth than landmarks with excellent AI-based predictions. Some landmarks, where AI-based prediction showed better accuracy, were omitted to increase the visibility of the plot



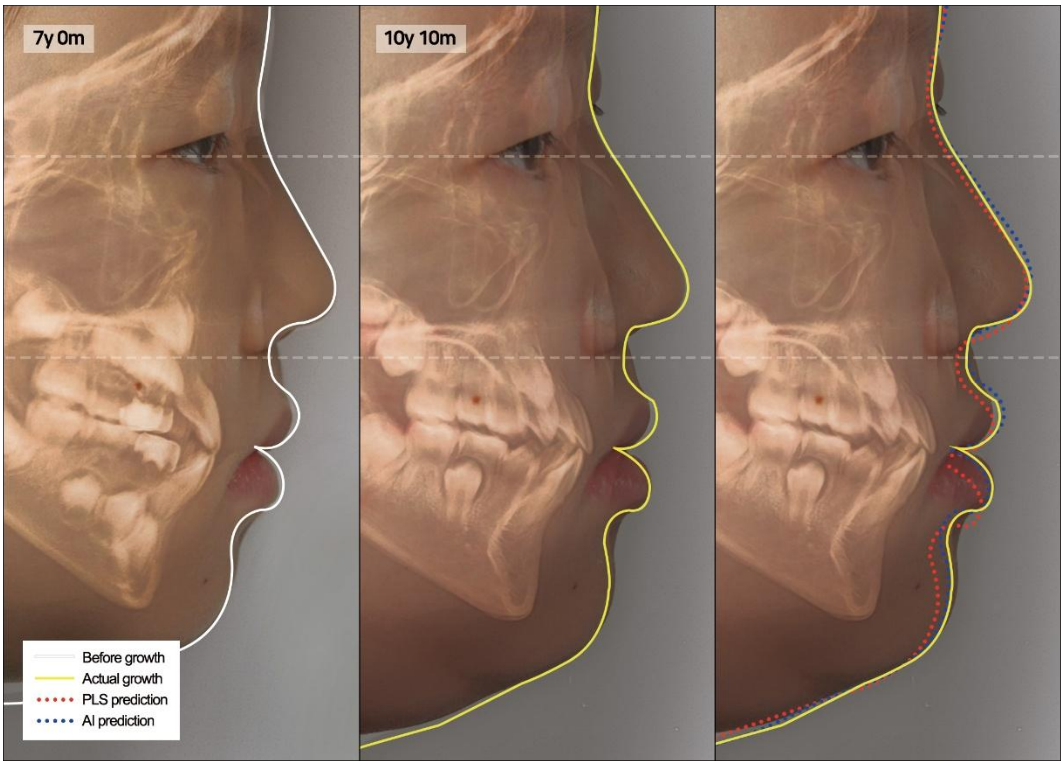


Figure 8, A. Comparisons between actual growth and prediction results in patient with Class I malocclusion. To concisely showcase the prediction result, only soft tissue outlines are shown.

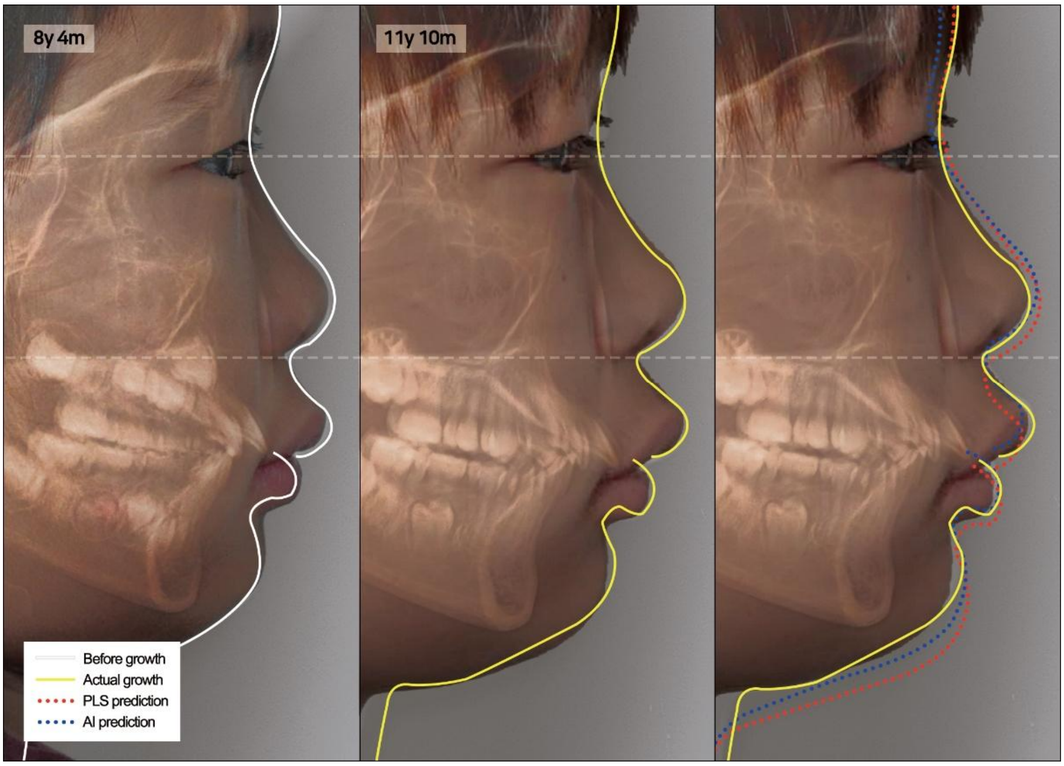


Figure 8, B. Comparisons between actual growth and prediction results in patient with Class II malocclusion. To concisely showcase the prediction result, only soft tissue outlines are shown.

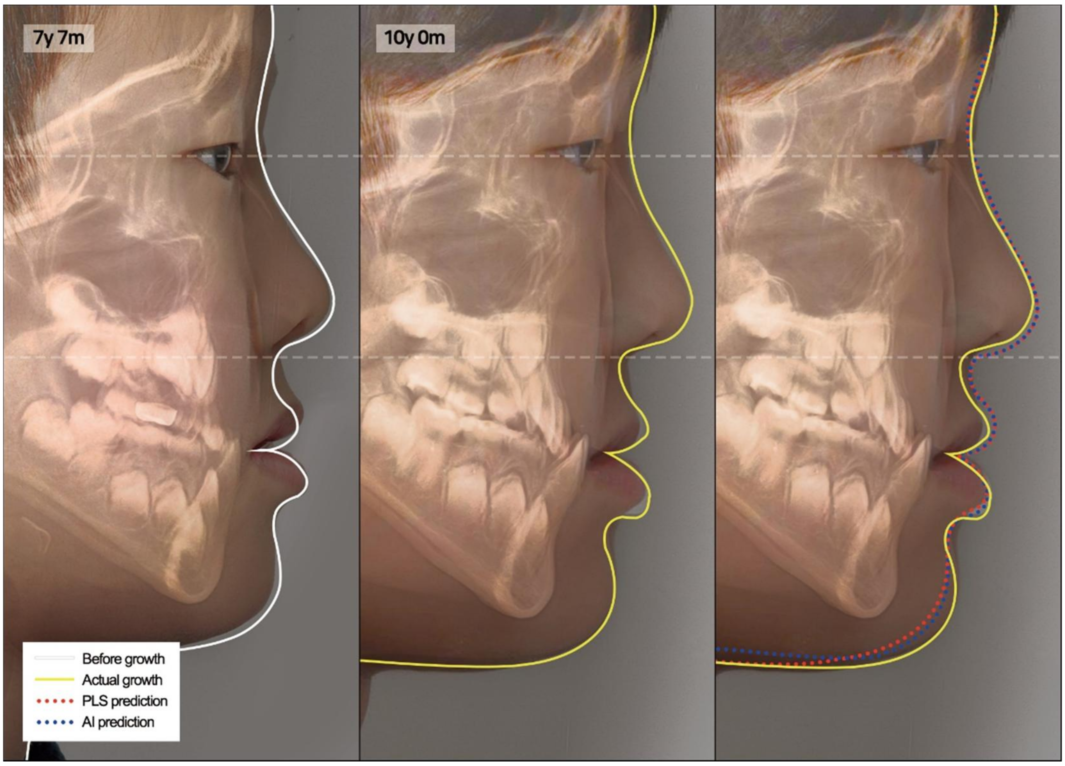


Figure 8, C. Comparisons between actual growth and prediction results in patient with Class III malocclusion. To concisely showcase the prediction result, only soft tissue outlines are shown.

국문초록

부분최소제곱 방법과 인공지능을 이용한 측모 성장 예측 방법 개발 및 비교 연구

문 준 호

서울대학교 대학원 치의과학과 치과교정학 전공

(지도교수: 이 신 재)

연구 목적: 두개안면 성장은 오랫동안 임상 교정 영역에서 중요한 주제로 여겨져 왔다. 성장을 정확하게 예측하고자 다양한 방법들이 개발되었으나, 성장의 개인 간 변이가 이를 예측하는 것을 어렵게 만든다. 또한, 기존의 통계적인 기법을 사용하는 것의 한계는, 예측해야 할 변수와 예측에 사용되는 변수들의 수가 많으며 서로 상관관계가 높다는 것이다.

부분최소제곱 (Partial least squares, PLS) 방법은 변수가 많고 상관관계가 높은 경우에도 통계적인 예측에 사용할 수 있으며, 양악수술 후 연조직 변화를 기존의 최소제곱 (Ordinary least squares, OLS) 방법보다 더 정확하게 예측할 수 있는 것으로 밝혀졌다. 이러한 PLS 방법은 개인의 성장 차이를 고려하고 상관관계 문제를 해결함으로써 성장을 예측하는 데 유용할 수 있다.

한편, 치과 영역에서 인공지능 (Artificial Intelligence, AI)의 활용이 주목받고 있다. 최근에 개발된 TabNet 은 도표 데이터에 심층 신경망 (Deep neural network, DNN)을 적용하기 위해 개발되었다. 도표로 정리된 성장 데이터에 TabNet 알고리즘을 적용한다면 정확한 성장 예측이 가능할 수도 있다.

본 연구의 목적은 개인의 골격과 연조직 특성을 반영한 성장 예측 모형을 개발하는 것이다. 이를 위하여 PLS 방법과 AI 를 이용한다. 그리고 두 가지 예측 모형의 성능을 평가하고 비교하는 것이다.

재료 및 방법: 교정 치료를 받은 적이 없는 303 명의 성장 환자 (여아 166 명, 남아 137 명)의 연속된 측면 두부계측방사선사진을 연구 대상으로 하였다. 다변량 PLS 알고리즘과 TabNet 심층 신경망 AI 알고리즘을 적용하여 성장 예측 모형을 개발하였으며, 예측 변수는 161 개였다. 반응 변수는 78 개의 두부계측점으로 구성되었다. 두 예측 모형의 정확성을 비교하기 위해 T-검정이 시행되었다. 성장 예측 오차에 영향을 미치는 요인을 조사하기 위해 다중 선형 회귀 분석이 시행되었다. 예측 오차의 패턴을 조사하고 예측의 정확성에 대한 성장 변동성의 영향을 평가하기 위해 신뢰 타원을 도해하였다.
결과: Leave-one-out 교차 검증을 사용하여 30 개의 PLS 성분을 가진 PLS 모형을 개발하였다. AI 기반 예측 모형의 경우, 하이퍼파라미터 튜닝 후 최적의 하이퍼파라미터가 선택되었다. 78 개의 계측점 중 55 개의 계측점에서 AI 기반 모형이 더 정확했다. PLS 방법은 두개저의 계측점을 포함한 10 개의 계측점에서 더 정확했는데, 이러한 계측점에서는 성장 변동성이 적었다.

나머지 13 개의 계측점에서는 두 방법 사이에 통계적 차이가 관찰되지 않았다. 불확실성이 높을 때는 성장 예측에 AI 를 활용하는 것이 더 유리했다. 평균적으로 AI 기반 모형은 PLS 기반 모형보다 1.11 mm 적은 예측 오차를 보였다. 예측을 시행하는 나이가 낮을수록 예측 오차가 더 컸다 (연간 0.01 mm). 또한 성장 예측 기간에 비례하여 예측 오차가 증가하였다 (연간 0.14 mm). 여자아이, II 급 부정교합, 골격 계측점, 상악골 계측점의 예측 결과는 남자아이, I 급 또는 III 급 부정교합, 연조직 계측점, 하악골 계측점 각각에 비해 더 정확했다.

결론: AI 를 이용한 성장 예측은 PLS 를 이용한 예측보다 더 정확했다. 이때, 예측 모형의 예측 오차는 성장 잠재력에 비례했다. PLS 와 AI 를 이용한 성장 예측은 개인의 특성을 반영하고 많은 수의 계측점을 예측하기 위해 다수의 예측 변수를 포함시킬 수 있는 접근 방식으로 보인다.

주요어: 성장 예측, 부분최소제곱 방법, 인공지능, 딥 러닝

학번: 2018-38862